Uncertainty in economic evaluations

Mohseninejad, Leyla

Uncertainty in economic evaluations: implications for healthcare decisions

Leyla Mohseninejad

Financial support for this study was provided by the National Institute for Public Health and the Environment (RIVM) as well as a grant from The Netherlands Organization for Health Research and Development (ZonMW).
Printing of this thesis was partially supported by the University of Groningen, the Research Institute SHARE of the Graduate School of Medical Sciences, and the University Medical Center Groningen.

ISBN

978-90-367-6632-6

Printing

Printed by  Lovebird design & printing solutions.
www.lovebird-design.com

Cover

Cover design by Leyla Mohseninejad.

RIJKSUNIVERSITEIT GRONINGEN

Uncertainty in economic evaluations: implications for healthcare decisions

Proefschrift

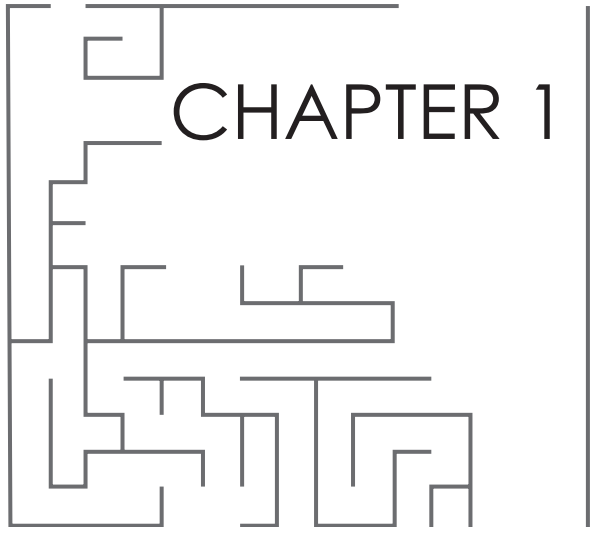
ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
woensdag 18 december 2013
om 12:45 uur

door

Leyla Mohseninejad
geboren op 14 november 1984
te Mashhad, Iran

Promotor:	Prof. dr. E. Buskens
Copromotor:	Dr. T.L. Feenstra
Beoordelingscommissie:	Prof. dr. J.C. Kluin-Neelemans
	Prof. dr. M.J. Postma
	Prof. dr. C.A. Uyl-de Groot

TABLE OF CONTENTS



General Introduction

New medications, treatments and technological possibilities are developed everyday all over the world to improve healthcare services. Before entering the market, new technologies pass filters of safety and efficacy. These criteria are essential, but they are not all that matter. Limited healthcare budgets impose a third important criterion: cost effectiveness.

Economic evaluations, especially cost-effectiveness analyses, play an increasing role in supporting policy making. In these methods, additional costs and health benefits of a new drug/health intervention are compared to the standard of care. When the additional costs and benefits of the new technology are balanced in a way that the adoption is worthwhile, the health policy can change in favor of the new technology. However, it is not always straight forward to compute additional costs and health gains from implementing a new drug/intervention. Different diseases with various specifications and progression stages would need different ways of calculating the long term costs and health effects. To inform the calculation of costs and effects, decision analytic models have been developed as a tool to support cost-effectiveness analysis for different conditions. These models help to synthesize all the characteristics of the disease and all the available evidence and find the costs and effects of the alternative interventions of interest.

After finding the best way to model a disease, cost-effectiveness analysis (CEA) helps decision makers to choose among different dugs/treatment options for a certain disease. In such analysis, costs are related to a single, common effect that may differ in magnitude among the alternative drugs/treatments (Drummond et al., 2005). The results of cost-effectiveness analysis are often incorporated in new guidelines, supporting the decisions makers to find the best way of allocating resources.

However, since any assessment of the effects as well as costs will remain uncertain to some degree, any decision based on cost-effectives will also be uncertain. This uncertainty arises from different sources, and can have an important effect on the results.

It should be noted that uncertainty is different from variability and patient heterogeneity. Variability refers to natural variation in quantities due to chance (e.g. when the effect of a certain drug is varying among patients with the same

characteristics) and patient heterogeneity refers to variations due to different characteristics of patients (e.g. when the drug effect is varying among patients with different characteristics). In contrast, uncertainty refers to the degree of precision with which quantities are measured (for instance when the effects on a particular patient is not precisely known) (van Belle, 2002). Uncertainty can often be reduced by gathering further information, while variability and patient heterogeneity are inevitable parts of the results of any economic evaluation. The variability between subjects has also been referred to as "first order uncertainty" or "stochastic uncertainty" in medical decision making literature (Groot Koerkamp, 2009; Stinnett and Paltiel, 1997).

In this thesis I aim to analyze the methods to handle the uncertainty in the results of economic evaluations, while I also address the variability and heterogeneity. There are three reasons why it is essential to analyze the uncertainty surrounding effects and costs (Claxton, 2008): (i) to evaluate the expected effects and costs correctly; (ii) to make sure if the current evidence is sufficient; and (iii) to consider the possible consequences of an uncertain decision.

In the next sections, different decision modeling approaches are introduced and some basics regarding the cost-effectiveness analysis are explained. Then different possible sources of uncertainty are discussed and some methods to handle the uncertainty are reviewed. In the end, the theme and objectives of this thesis are described.

DECISION ANALYTIC MODELING

Decision analytic models use mathematical relationships to find possible consequences of different alternative decisions (Briggs et al., 2006). In a decision analysis, inputs related to each decision option are fed into to the model. Using the inputs, the model will generate the likelihood of each outcome together with costs and health effects.

An important aspect of decision modeling, which is also of interest for this thesis, is to allow for the uncertainty and variability in the outcomes. The inputs are uncertain and hence the outcomes are uncertain, and the decision models make it

possible to analyze the uncertainty and to quantify it. Different model structures are used in this thesis in order to calculate the costs and effects and the uncertainties. Based on the features of the disease and the technology that is being evaluated, different model structures might be appropriate (Briggs et al., 2006). In the next sections, some of the most important model structures are described.

Decision Trees

The decision trees are probably the simplest and the most common form of decision models (Briggs et al., 2006; Drummond et al., 2005). They illustrate a patient's possible pathways, including possible prognosis tests and interventions. By calculating the probability of each pathway together with costs and effects of that pathway, and summing up the costs and effects weighed by the probabilities across the pathways, the expected costs and effects can be computed (Drummond et al., 2005). An example of a decision tree model is given in chapter 2 of this thesis (Mohseninejad et al., 2012).

Decision trees are widely applied in economic evaluations, but they have some important limitations. First, since there is no explicit time element in the decision trees, any decision which needs to include the time aspect would be hard to model. Second, decision trees can get very complicated when modeling long term, and in particular, chronic diseases. (Drummond et al., 2005).

Markov Models

Markov models have overcome the limitations of decision trees and they are widely used in economic evaluations (Drummond et al., 2005). Markov model explicitly model time, and are suitable to present the progression of chronic diseases. In a Markov model, the disease cycle is divided into a number of distinct states. By having the transitions probabilities for movement between these states and attaching estimates of costs and health effects to each state, the long term costs and effects associated with a particular healthcare intervention/medicine can be estimated (Briggs and Sculpher, 1998). Although the Markov models have been very efficient in handling economic evaluations, they also have some restrictions.

The important restriction is related to the Markov model's assumption of being "memoryless". This assumption means that when a patient enters a state, the model will not consider where the patient has come from or how long the patient has been in the process. However, some extensions of the Markov models have been designed which incorporate time dependency to the models (Briggs et al., 2006). In the third chapter of this thesis, we show this time dependency in a Markov model for a depression prevention program (Mohseninejad et al., 2013).

Patient Level Simulation

In a patient level simulation, the patients are tracked individually while they are moving through the model and experiencing particular events. Hence, instead of transition probabilities, the interest is on the time of the next event for each particular patient (Karnon, 2003). Patient level simulation models (also called as microsimulations or individual sampling methods) offer great flexibility, as they are able to account for the history of the patients. Despite the structural flexibility, such models also have some limitations. Since the parameters representing possible future pathways for a patient are conditional on history, additional evidence is needed to populate such models. The simulation requirement of these models can be sometimes very time consuming (Drummond et al., 2005).

In chapters 4,5 and 6 of this thesis, we use some sort of patient level simulations to accumulate costs and effects over a certain period of time and find final estimations.

COST-EFFECTIVENESS ANALYSIS

Cost effectiveness analysis (CEA) is generally referred to analyses in which costs are related to a single, common effect that might be different for alternative health programs/medicines (Drummond et al., 2005). Results of such comparison are mainly stated as costs per unit of effect.

Cost effectiveness ratio

Typically the CEA is expressed in terms of a ratio where the denominator is a gain in health (life years/ Quality adjusted life years, improved symptoms, etc) and the numerator is the cost associated with the health gain (Gold et al., 1996). To make the comparison of alternative options easier, incremental cost-effectiveness ratios (ICERs) have been introduced. If C_A and C_B are the costs of the health intervention/ medicine A and B respectively, and E_A and E_B are the effects, *ICER* of intervention A versus B will be computed as:

$$ICER = \frac{C_A - C_B}{E_A - E_B} \quad (1)$$

When the ICER is below the maximum acceptable cost-effectiveness ratio (or willingness-to-pay threshold per unit of effect), the program A is considered cost-effective. In other words, if we call the threshold ratio by λ, we should have:

$$\frac{\Delta C}{\Delta E} < \lambda \quad (2)$$

Incremental Net Monetary Benefits

ICERs are very informative and are widely used to express results of economic evaluations. However, since they are ratio estimators, they sometimes cause problems for standard statistical methods (Briggs and Fenn, 1998). By re-expressing equation 2 we have:

$$\lambda\Delta E - \Delta C > 0 \quad (3)$$

The increase in health effects (ΔE) multiplied by the amount the decision maker is willing to pay per unit of health (λ), less the increase in costs (ΔC) is called the Incremental Net Benefits (INB). Working with INBs is often more convenient for policy makers and clinicians, as the extra benefit of the new technology can be easily extrapolated to the population level by multiplying the benefit per patient by the total population. Besides, the confidence intervals for INBs can give a more

clear view of the risk to the policy makers (Groot Koerkamp, 2009). In chapters 3,5 and 6 of this thesis, we choose to work with INBs rather than ICREs and we go on to show that the uncertainty surrounding INB can be more easily analyzed by assigning distributions.

SOURCES OF UNCERTAINTY

The uncertainty surrounding the estimations of the cost-effectiveness of a particular intervention or drug can be caused by different sources, for example, uncertainty in the treatment/drug effects or costs, the type of model used, and the applicability or generalizability of the results to the decision-maker (Bojke et al., 2009). There are several ways to categorize the sources of uncertainty. In a more common categorization, the sources of uncertainty are described as: parameter uncertainty, methodological, and structural uncertainty (Briggs, 2000).

Parameter uncertainty

Decision analytic models need some input parameters to generate the outputs such as health gains and costs. A decision model might need various input parameters to be able to perform. For instance the epidemiological parameters such as incidence and prevalence, relative risks, duration of the disease stages and the treatment, and related costs are the most common input parameters of a decision model. Estimating the true value of all the different parameters is almost impossible, meaning that there will always be an uncertainty surrounding the inputs of the model. Methods to analyze the uncertainty regarding the model parameters have been well established (Briggs et al., 2012) and recorded in health economic guidelines (e.g. (NICE, 2004)). Some of these methods are explained later in this chapter.

Methodological uncertainty

The uncertainty in outcomes of an economic evaluation could be due to the methods underpinning the evaluation. One example of the methodological source of uncertainty is the perspective adopted, which defines the selection of costs and

health gains to include in an economic evaluation (Briggs and Gray, 1999). Choosing the right perspective is essential to avoid extra uncertainty in the outcomes.

Assigning the maximum acceptable value for each additional unit of health has also been a source of debate. Since choosing the threshold value might need to reflect disease characteristics, budgetary impacts, equity measures and other criteria, it can add to the uncertainty of the evaluation (Birch and Gafni, 2006; McCabe et al., 2008; Stolk et al., 2004).

Another choice regarding the methodology concerns the selection of discount rates. Different rate to discount costs and health benefits have been proposed so far, and there has been a debate whether equal discount rates should be considered for costs and effects (Brouwer et al., 2005; Claxton et al., 2006). Selecting the right instrument for valuation of health outcomes is also a source of uncertainty in the methodology (Briggs, 2000).

While the focus in the literature of uncertainty analysis has mostly been on parameter uncertainty, in this thesis we pay considerable attention to the methodological uncertainties. We analyze the uncertainty caused by choosing the perspective (chapter 3), the threshold value (chapters 2, 3, 4, 5 and 6), the discount rate (chapters 2 and 5) and the valuation of health outcomes (chapter 2). We show how important this source of uncertainty source can be in different examples. In the third chapter of this thesis, we show that methodological uncertainty in choosing a perspective for the analysis is causing significant changes in the economic evaluation and the priority settings (Mohseninejad et al., 2013). We discuss the related findings in chapter 7.

Structural uncertainty

Other types of uncertainty which cannot be classified under parameter or methodological uncertainty are usually called structural uncertainty (Bojke et al., 2006). Different types of simplifications and scientific judgments made when constructing and interpreting any model are examples of structural uncertainty (Bojke et al., 2009). Structural uncertainty might not always be distinguishable from other types of uncertainty, for instance sometimes changing a model parameter

might change the structure of the model as well. A common type of structural uncertainty is the choice of the alternative model specification. Often, scenarios are presented based on extreme assumptions that can be made (Claxton et al., 2004). We present the scenario analysis to analyze the structural uncertainty in the next chapters of this thesis. In chapter 2 we illustrate the importance of structural uncertainty by showing how cost-effectiveness results and budgetary impact might change when including different scenarios. In chapters 4 and 5 we discuss different scenarios when deciding on the right time for making a decision. I address possible improvements in chapter 7.

METHODS TO HANDLE UNCERTAINTY

Various studies have so far addressed the uncertainty problem and several methods have been developed to handle the uncertainty which is rising from different sources. In this section I introduce some common uncertainty analysis methods which are going to be applied and extended in this thesis.

Deterministic Sensitivity analysis

Deterministic sensitivity analysis (DSA) is probably the simplest way to analyze and present the uncertainty in the outcomes. In a DSA, parameter values are varied manually to examine the sensitivity of the model's results to specific parameters or sets of parameters (Briggs et al., 2012).

Probabilistic Sensitivity analysis

Probabilistic Sensitivity analysis (PSA), evaluates the joint effect of uncertainty about all estimated parameter values in the model. In PSA, a probability distribution is assigned to each parameter of the model. Then all parameters are varied simultaneously, with multiple sets of parameter values being sampled from the priori–defined distributions simultaneously using Monte Carlo simulation (Briggs and Gray, 1999; Briggs et al., 2012). It is important to assign appropriate distributions to different parameters with respect to the characteristics of the parameter being estimated (Briggs et al., 2006).

Different ways of presenting and analyzing the results of a probabilistic sensitivity analysis are described later in this chapter.

Cost effectiveness plane

The most common way of presenting the results of a probabilistic sensitivity analysis is the cost-effectiveness (CE) plane (Anderson et al., 1986; Black, 1990).

An illustration of the CE plane is presented in Figure 1. In the diagram, the horizontal access represents the gains in effects when using the new intervention (A) versus the old standard one (B), and the vertical access represents the additional costs of A comparing to B. The plane is divided into four quadrants indicating four possible situations in relation to the additional costs and additional health outcome of the new treatment compared to the old standard one. Results of probabilistic sensitivity analysis can be shown by dots in the plane: each dot represents the result of one run of the probabilistic model.

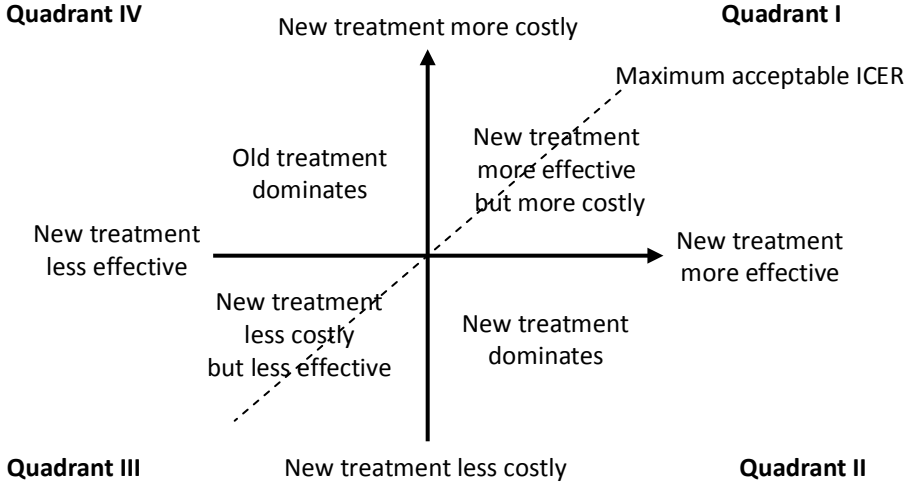


Figure 1 The general form of a cost-effectiveness plane. Adapted from Black (Black, 1990)

CE plane is very useful for the presentation of uncertainty as a region in the cost-effectiveness space (van Hout et al., 1994). For instance if the estimates are all distributed in the same quadrant of the plane, it can be concluded that the uncertainty is relatively low.

Cost-effectiveness acceptability curves

To consider the effect of different willingness-to-pay (λ) thresholds in the cost-effectiveness analysis, cost-effectiveness acceptability curves (CEACs) are introduced. For each predefined λ, CEACS show the probability that the cost-effectiveness ratio found in the study is acceptable (van Hout et al., 1994). The probabilities are calculated as the proportion of dots which fall below the λ line (maximum acceptable ICER) in figure 1.

Value of information analysis

With roots in statistical decision theory (Raiffa and Schlaifer, 1959), Value of information (VOI) analysis has been applied in the health care field in recent years (Claxton, 1999; Claxton and Posnett, 1996). VOI analysis can be performed as a part of probabilistic sensitivity analysis to show if additional information is needed to support the decision on adoption or reimbursement of a health intervention/ drug. Such analysis can also support the decisions to choose among different types of research (Claxton and Sculpher, 2006), select the parameters which are more important in future research (Ades et al., 2004), or to find the optimal design of studies (Claxton and Posnett, 1996; Claxton and Thompson, 2001; Eckermann and Willan, 2007).

In this thesis we develop rather novel applications of the value of information analysis to show the robustness of the results of economic evaluations (chapters 2 and 3), to establish the priorities in future research with respect to perspective (chapter 3), to investigate the information update over time (chapters 4, 5 and 6), and to evaluate quality of data (chapter 6).

Resolving uncertainty over time

Various methods have so far been introduced to address the time aspect in decision uncertainty. As the new technologies/drugs spread over time, some part of uncertainties on their costs and effectiveness resolves; hence it is important to investigate the effect of time in the information flow which leads to decreased uncertainty. Many different methods have been developed or adapted from other

fields of science to healthcare, aiming to analyze the resolution of uncertainty over time and the subsequent decisions. One of the more common methods is the applications of real options approach (ROA) (Raiffa and Schlaifer, 1959) in health technology assessment (Palmer and Smith, 2000). Some extensions of the value of information analysis in optimal trial design (Willan and Pinto, 2006) have also been used to address the information update over time (Chen and Willan, 2013; Willan and Kowgier, 2008). Sequential analysis (Wald, 1945) is another related method for deciding on adoption/reimbursement.

Although many approaches have been developed to analyze the information trend over time, there is still the need to find the optimal time at which delaying a decision further to wait for more information is no longer worthwhile. In this thesis we discuss various methods towards analysis of uncertainty over time (chapters 4 and 5), we develop a new method to find the optimal time of decision making (chapter 4), and we show applications of our method (chapters 5 and 6)

POLICIES TOWARDS UNCERTAINTY MANAGEMENT

When a new healthcare technology is introduced the information about its cost-effectiveness is often scarce. Hence, it is essential to analyze the uncertainty before making a decision. In some instances a cost-effectiveness evaluation may be a onetime event, used when making the adoption/reimbursement decision for a new drug/technology. For such cases, analysis of uncertainty would enable the decision maker to get knowledge on the risk in the decision and the consequences of a wrong choice before actually making the decision. It will also inform the decision maker whether further research is needed before making a choice, and if so, what sort of research is needed. We address the policy implications regarding the uncertainty analysis for onetime evaluations in chapter 2 and 3 of this thesis.

In many other cases, the uncertainty surrounding the information about a new drug/technology is so high that the decision maker cannot decide whether to adopt/reimburse right away, needing some time to observe evidence about the costs and effects of the new technology. On the other hand, providers and patients legitimately require access to the new technology from which they might

benefit. These two rather opposite demands have emerged new policy schemes for managing the uncertainty surrounding the cost-effectiveness of new health technologies in recent years. Different coverage and reimbursement schemes have been proposed to guarantee further research before making a final decision in different countries (e.g. (Stafinski et al., 2010; Stafinski et al., 2011; Walker et al., 2012)). Access with evidence development (AED) is a category of these schemes (Carlson et al., 2010). AED refers to a form of a provisional coverage arrangement in which the new drug/technology is temporarily funded until further evidence supports a definite adoption/reimbursement decision.

We explore different policy options and contribute to the methods of linking administering and reimbursement of medical products to the collection of additional evidence in an AED scheme in chapters 4, 5, 6 and 7 of this thesis.

AIM AND SCOPE OF THIS THESIS

The objective of this thesis is to explore different ways of handling the uncertainty in economic evaluation of new medical technologies and contribute to the methods of the uncertainty analysis.

After the general introduction on decision analytic modeling, cost-effectiveness analysis and analysis of uncertainty in **chapter 1**, an elaborated example of the decision modeling followed by cost-effectiveness and uncertainty analysis is presented in **chapter 2**. In this chapter, a decision tree is used to model the cost-effectiveness of screening for coeliac disease in patients with irritable bowel syndrome (IBS). A deterministic sensitivity analysis is performed to examine the effect of different model parameters in the results. Also probabilistic sensitivity analysis including the presentation of CEACs and value of information analysis is used to analyze the uncertainty in the outcomes. Results show that the screening program is cost-effective, and that uncertainty surrounding the outcomes have limited effect on the decision.

In **chapter 3** a Markov model is used to evaluate the costs and long term health benefits of screening followed by Minimal Contact Psychotherapy (MCP) for depression prevention. In this chapter, value of information analysis is performed

and presented in a more elaborated way and results are compared from a healthcare and a societal perspective. Results of this chapter show the need for carefully choosing the relevant perspective for the decision problems.

Chapter 4 goes one step further in handling the uncertainty in economic evaluations by including timing in the reimbursement decisions. In this chapter we use a value of information framework to develop a model which describes the problem of making a definite decision on a conditionally reimbursed drug. The method developed in this chapter enables the decision maker to select the optimal period of conditional reimbursement and additional evidence gathering. By choosing the optimal time for a definite decision, more flexibility is introduced into the methods of solving the uncertainty over time.

In **chapter 5** simulation methods are used to find the optimal time of making a definite decision on reimbursement of voriconazole for primary treatment of invasive aspergillosis. While the old regulations for expensive new inpatient drugs in The Netherlands specify a maximum period of four years for conditional funding, it is shown in this chapter that in case of antifungal drugs at any point after the 2nd year the uncertainty becomes almost ignorable and the decision is close to optimum.

Chapter 6 evaluates a registry which has been set up to provide further evidence for deciding on reimbursement of oxaliplation for treatment of stage III colon cancer. In this chapter, patient level data has been mixed with simulated data to find out if the registry provides sufficient information to resolve the uncertainty in the reimbursement decision of oxaliplatin. Results of this chapter indicate that the registry should have been improved, or should have been stopped after 2 years rather than simply complete the current 4-years period.

The concluding **chapter 7** presents a discussion on methods used in this thesis, the methodological contribution of the thesis and the policy implications.

REFERENCES

Ades A, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. *Medical Decision Making* 2004; 24(2): 207-227.

Anderson JP, Bush JW, Chen M, Dolenc D. Policy space areas and properties of benefit–cost/utility analysis. *The Journal of the American Medical Association* 1986; 255(6): 794-795.

Birch S, Gafni A. Information created to evade reality (ICER): things we should not look to for answers. *Pharmacoeconomics* 2006; 24(11): 1121-1131.

Black WC. The CE plane: A graphic representation of cost-effectiveness. *Medical Decision Making* 1990; 10: 212-214.

Bojke L, Claxton K, Palmer S, Sculpher M. Defining and characterising structural uncertainty in decision analytic models. CHE Research Paper 9. Centre for Health Economics, University of York:2006; York.

Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing Structural Uncertainty in Decision Analytic Models:A Review and Application of Methods. *Value in Health* 2009; 12: 739–749.

Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation*. Oxford University Press: Oxford 2006.

Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Economics* 1998; 7(8): 723-740.

Briggs A, Gray A. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technology Assessment* 1999; 3(2): 1-134.

Briggs A, Sculpher M. An Introduction to Markov Modelling for Economic Evaluation. *Pharmacoeconomics* 1998; 13(4): 397-409.

Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000; 17(5): 479-500.

Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty Analysis: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. *Medical Decision Making* 2012; 32: 722–732.

Brouwer W, Niessen L, Postma M, Rutten F. Need for differential discounting of costs and health effects in cost effectiveness analyses. *British Medical Journal* 2005; 331(7514): 446-448.

Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy* 2010; 96: 179–190.

Chen MH, Willan AR. Determining optimal sample sizes for multistage adaptive randomized clinical trials from an industry perspective using value of information methods. *Clin Trials* 2013; 10: 54–62.

Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 1999; 18(3): 341-364.

Claxton K. Exploring Uncertainty in Cost-Effectiveness Analysis. *Pharmacoeconomics* 2008; 26(9): 781-798.

Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S. A pilot study on the use of decision theory and value of information analysis as part of the NHS health technology assessment programme. *Health Technology Assessment* 2004; 8(31): 1-103.

Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Economics* 1996; 5(6): 513-524.

Claxton K, Sculpher M, Culyer A, McCabe C, Briggs A, Akehurst R, Buxton M, Brazier J. Discounting and cost-effectiveness in NICE - stepping back to sort out a confusion. *Health Economics* 2006; 15(1): 1-4.

Claxton K, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics* 2006; 24(11): 1055-1068.

Claxton K, Thompson K. A dynamic programming approach to the efficient design of clinical trials. *Journal of Health Economics* 2001; 20(5): 797-822.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press: 2005.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. *Health Econ* 2007; 16: 195–209.

Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine*. Oxford University Press: Oxford;1996.

Groot Koerkamp B. Uncertainty in medical decision making, Knowing how little you konw. PhD Thesis, Erasmus MC. Rotterdam: 2009.

Karnon J. Alternative decisionmodelling techniques for the evaluation of health care technologies:Markov processes versus discrete event simulation. *Health Economics* 2003; 12: 837–848.

McCabe C, Claxton K, Culyer A. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 2008; 26(9): 733-744.

Mohseninejad L, Feenstra T, van der Horst HE, Woutersen-Koch H, Buskens E. Targeted screening for Coeliac Disease among irritable bowel syndrome patients: analysis of cost-effectiveness and value of information. *European Journal of Health Economics* 2012; November 2012: [Epub ahead of print].

Mohseninejad L, van Baal PHM, van den Berg M, Buskens E, Feenstra T. Value of Information Analysis from a Societal Perspective: A Case Study in Prevention of Major Depression. *Value in Health* 2013: In Press.

NICE. Guide to the Methods of Technology Assessment. National Institute for Clinical Excellence (NICE):2004; London.

Palmer S, Smith PC. Incorporating option values into the economic evaluation of health care technologies. *J Health Econ* 2000; 19: 755-766.

Raiffa H, Schlaifer R. *Probability and statistics for business decisions*. McGraw-Hill: New York;1959.

Stafinski T, McCabe C, Menon D. Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *Pharmacoeconomics* 2010; 28(2): 113-142.

Stafinski T, Menon D, Davis C, McCabe C. Role of centralized review processes for making reimbursement decisions on new health technologies in Europe. *Journal of ClinicoEconomics and Outcomes Research* 2011; 3: 117-186.

Stinnett A, Paltiel A. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. *Medical Decision Making* 1997; 17(4): 483-489.

Stolk E, van Donselaar G, Brouwer W, Busschbach J. Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional shortfall. *Pharmacoeconomics* 2004; 22(17): 1097-1107.

van Belle G. *Statistical Rules Of Thumb*. Wiley-Interscience: 2002.

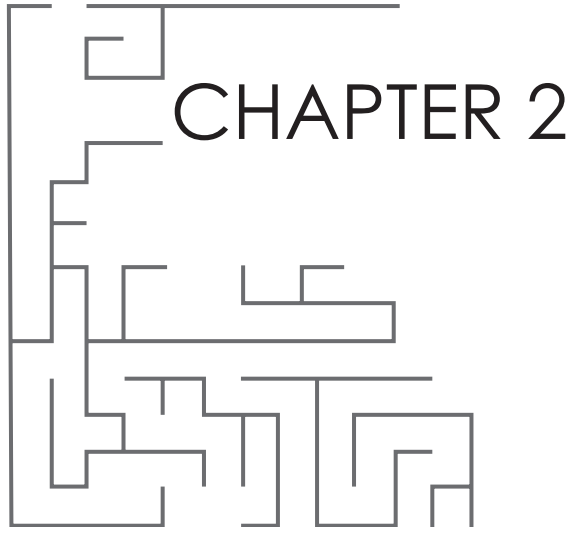
van Hout B, Al M, Gordon G, Rutten F. Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics* 1994; 3: 309–319.

Wald A. Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics* 1945; 16(2): 117-186.

Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing or patient access scheme? A framework for coverage decisions. CHE Research Paper 77. Centre for Health Economics, University of York:2012; York.

Willan A, Kowgier M. Determining optimal sample sizes for multi-stage randomized clinical trials using value of information methods. *Clin Trials* 2008; 5(4): 289–300.

Willan A, Pinto E. The expected value of information and optimal clinical trial design. *Stat Med* 2006; 24(21): 1791–1806.



Targeted screening for Coeliac Disease among Irritable Bowel Syndrome patients: Analysis of cost-effectiveness and value of information.

Leyla Mohseninejad

Talitha Feenstra

Henriëtte E. van der Horst

Hèlen Woutersen-Koch

Erik Buskens

European Journal of Health Economics. 2012 [Epub ahead of print]

Abstract

Objectives: A high prevalence of Coeliac Disease (CD) is found among patients with a clinical diagnosis of irritable bowel syndrome (IBS) compared to the general population. Symptoms of CD are quite similar to IBS, but its treatment is different. The aim of this study is to evaluate the cost-effectiveness of screening for CD in patients with diarrhoea/ mixed type IBS (IBS-D/mix) in terms of cost per QALY in the Netherlands.

Methods: A decision model was constructed to evaluate the costs and health benefits of serological testing followed by confirmatory endoscopy with biopsy. Probabilistic sensitivity analysis (PSA) was performed to examine the effect of parameter uncertainty. Finally, the budget impact of implementing the screening process was also computed for implementation over a 10-year time horizon.

Results: Screening resulted in an increase of about 0.07 QALYs per patient over a lifetime horizon. The incremental cost effectiveness ratio was about 6,200 €/QALY compared to no screening. PSA showed that the uncertainty in cost effectiveness results is not considerable. Value of information analysis confirmed the robustness of the results. Screening all current patients with diarrhea/mixed type IBS would require a total budget of about 25 million Euros over a 10 year time period.

Conclusion: Screening patients with IBS-D or IBS-mix for CD is almost certainly cost-effective. The screening program would improve the quality of life of those patients with IBS symptoms who actually have CD at a relatively low cost.

INTRODUCTION

Irritable bowel syndrome (IBS) is a functional gastrointestinal disorder characterized by abdominal pain and defecation pattern disturbances (Inadomi et al., 2003). IBS is quite common, with a prevalence ranging from 6.2% to 12% in Europe (Hungin et al., 2003). Compared to the general population patients suffering from IBS have a low quality of life. In recent years, several quality of life instruments have been applied to IBS populations. Generic questionnaires that result in values useful for computing quality adjusted life years (QALYs) have been validated against more disease specific and multidimensional questionnaires. The EQ5D turned out to be valid with scores ranging from 0.62 to 0.75 in different IBS populations (Akehurst et al., 2002; Berg et al., 2006; Bracco et al., 2007; Brazier et al., 2006; Brazier et al., 2002; Brazier and Usherwood, 1998; Bushnell et al., 2006; Spiegel et al., 2009). IBS shows an association with Coeliac Disease (CD), a chronic autoimmune disorder with a severe impact of symptoms on quality of life (Gray and Papanicolas, 2010; Sanders et al., 2001; Sanders et al., 2003; Shahbazkhani et al., 2003) and increased mortality risk. Once detected, CD should be treated by prescribing a gluten free diet. For this reason, diagnosis of CD in IBS patients is important. Normally, such diagnosis may occur with a substantial delay, due to the similarity of symptoms.

While screening for CD in the general population has been shown not to be cost-effective (Hershcovici et al., 2010), conducting screening for CD in patients with IBS has been suggested to be potentially worthwhile. Studies from Mein and Ladabaum (Mein and Ladabaum, 2004) and the National Institute for Health and Clinical Excellence (NICE) guideline on IBS (2008a) have assessed the cost-effectiveness of screening for CD in all patients with IBS symptoms. However, constipation only is not among symptoms of CD (Spiegel et al., 2004). Therefore, we can assume that the prevalence of CD in IBS patients with only constipation type symptoms is low and comparable to that in the general population. Hence, excluding patients with constipation from the screening process and including only those with diarrhoea predominant or mixed type IBS (both constipation and diarrhoea) may result in a more favourable cost effectiveness ratio. Spiegel et al. (Spiegel et al., 2004) have evaluated testing for CD in IBS patients with predominant diarrhoea. They found

that such a targeted screening program has an acceptable cost per symptom improvement when the prevalence of CD is above 1%. The health gains ensuing from a correct diagnosis of CD were modelled quite different in the three studies mentioned above. Mein and Ladabaum (Mein and Ladabaum, 2004) only took into account utility gains, while the research for the NICE guideline (2008a) considered only survival gains. Spiegel et al. (Spiegel et al., 2004) considered symptomatic improvements as the outcome measurement.

That is, no study evaluated costs per QALY for targeted screening taking into account both health benefits in terms of survival and in terms of quality of life.

The aim of the current study hence was to evaluate the cost-effectiveness of targeted screening for CD in patients with IBS-D/ IBS-mix in the Netherlands in comparison to no screening and to screening all IBS patients, in terms of cost per QALY. We took into account both the improvements in quality of life and reductions in mortality resulting from adhering to a gluten free diet and used QALYs as the outcome measurement. We reported the number of cases that may be detected by the screening process and identified the most important sources of uncertainty in the evaluation. The effect of parameter uncertainty was evaluated by means of assigning distributions to parameters and running probabilistic sensitivity analysis. In addition, elaborate univariate sensitivity analyses were performed to test the effect of several model assumptions. Cost-effectiveness acceptability curves were derived and the budgetary impact of introducing screening for CD in the Dutch population was also assessed. In addition, a value of information analysis was performed to assess the need for additional research on uncertain parameter values.

METHODS

A decision model was constructed to compare the costs and health benefits of screening and subsequent treatment for Coeliac Disease (CD) among patients with IBS-D/mix with care as usual that is, no structured screening strategy. The model reflects possible trajectories over the life course of a cohort of IBS patients.

In the model, undiagnosed CD patients have an increased mortality, as well as reduced quality of life compared to diagnosed CD patients. However, in order to

avoid overestimating the health gains of screening, cases of CD among IBS patients are assumed to be detected because of sustained symptoms after a delay. The duration of the delay period is hard to establish. In previous studies delays that were applied ranged from 6 months to infinity (no detection of CD in case of no screening). We used an estimate of four years for the delay period. Such estimate is quite conservative, since evidence from the United States and Europe indicates that most CD patients probably remain undiagnosed for the rest of their lives (Spiegel et al., 2004). We included smaller and larger delay times in sensitivity analysis to examine the effects of the different delay time assumptions.

Patient population

We divided bowel habits of IBS patients into three categories: diarrhoea, constipation and mixed (van der Veek et al., 2007). The prevalence of CD among those IBS patients whose bowel habits are restricted to constipation is assumed to be similar to the CD prevalence in general population. Accordingly we excluded patients with constipation in the base case analysis. However, they have been included in sensitivity analysis to check whether such exclusion significantly improved the cost effectiveness of the testing strategy. Cohort age at screening was assumed to be 34, based on the expert opinions about typical IBS patients.

Screening strategy

Screening starts with tissue transglutaminase antibody (tTG) in combination with Immunoglobulin A (IgA) antibody test according to the current Dutch practice. When the IgA level is lower than 0.7 g/l, the results of the tTG are not reliable and endoscopy with biopsy will be conducted to confirm possible CD. When IgA level is higher than 0.7, tTG results are reliable and decision will be made based on them: patients with positive tTG results will enter endoscopy with biopsy and those with negative results will be assumed to have no CD. There is a small probability of major complication when endoscopy is performed (0.2%). Such complications increase costs and may be fatal in some cases (5%) (Mein and Ladabaum, 2004).

Model

A decision model was built using TreeAge software (TreeAge Pro 2009 Suite, Release 1.0.1). The control strategy is "no testing", in which we distinguished two possible outcomes. First, showing IBS symptoms which are not due to CD. In this case, the patient is a normal IBS patient, will receive routine IBS care and her annual costs and quality of life is computed based on published data from studies in IBS populations. Second, a patient with IBS symptoms may actually have CD. In this case, we assume that the CD will be detected after a delay period of four years. Therefore, in the first four years the patient will be considered to have the same costs as an IBS patient, but the quality of life and mortality of a CD patient. After four years, the costs of diagnosis of CD are incurred and from then on, the specific costs of IBS care will be replaced by the costs of a gluten-free diet and CD care, while quality of life improves and mortality decreases.

The intervention strategy is screening patients with IBS-D and IBS-mix. Here again, several possible outcomes exist. Patients may have IBS and test negative (true negatives). They may have IBS and test positive (false positives), in which case the endoscopy will show negative results. They may have CD and test negative (false negatives), in which case CD will only be detected after the delay period. Finally they may have CD and test positive (true positives), and receive a proper diagnosis of CD after endoscopy, followed by CD treatment.

Net present values for all future costs and health benefits have been calculated, applying a discount rate of 1.5% to health benefits and 4% to costs (2006). The decision model is illustrated in Figure 1.

The prevalence of CD among patients with IBS symptoms regardless of symptom types was 4.7% (2008b). However, excluding patients with constipation will lead to an increased prevalence in the screened population. About 25% of IBS patients have solely constipation symptoms (van der Veek et al., 2007), and using the 0.35% prevalence of CD in the general population (2008b) for them, the prevalence of CD among IBS patients eligible for screening should be 6.15%. Prevalence and incidence rates of diagnosed IBS were used to estimate the size of the target population for screening. The input parameters of the model, their base case values and the distributions used in the probabilistic sensitivity analysis are listed in table 1.

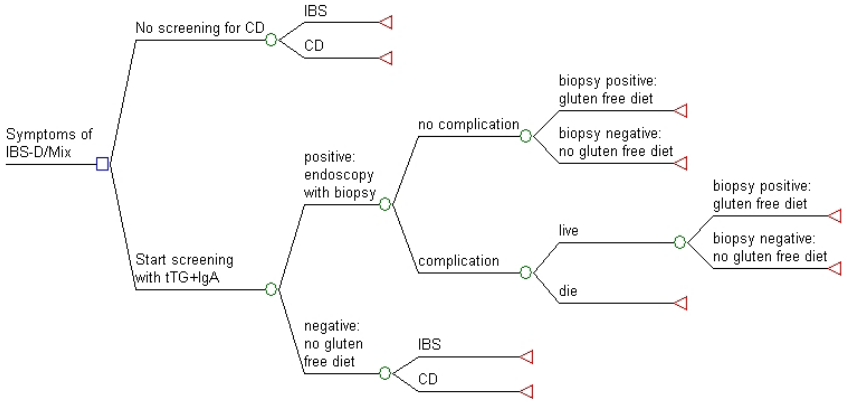


Figure 1 The decision model of screening for coeliac disease in patients with symptoms of IBS D/mix

Table 1 input parameters and distributions

Parameter	Value	Distribution	Reference
Clinical parameters			
Age	34 years	---	---
Prevalence of CD among patients with IBS	4.7 %	Beta (35,714)	(2008b)
Prevalence of CD in general population	0.35 %	Beta (6,1700)	(2008b)
Prevalence of constipation among patients with IBS	25%	Beta (13,39)	(van der Veek et al., 2007)
Prevalence of diagnosed IBS in general population	1.05%		(van der Linden et al., 2004)
Annual incidence of IBS	0.56%		(van der Linden et al., 2004)
IBS utility	0.675	Lognormal (-0.39,0.029)	(Akehurst et al., 2002)
CD utility	0.56	Lognormal (-0.58,0.026)	(Gray and Papanicolas, 2010)
Treated CD utility	0.84	Lognormal (-0.17,0.01)	(Gray and Papanicolas, 2010)
Standard Mortality Rate for untreated CD	1.6	---	(Shamir et al., 2006)
Test performance			
Sensitivity of tTG	0.949	Beta (26,1.40)	(2008b)
Specificity of tTG	0.975	Beta (106,2.71)	(2008b)
Probability of major complication with endoscopy	0.2 %	Beta (3.3,1600)	(Mein and Ladabaum, 2004)
Probability of death due to major complication	5 %	Beta (5.12,88)	(Mein and Ladabaum, 2004)
Costs (price level 2009)			
Cost of tTG	€ 39.63	Uniform (20-55)	(1998)
Cost of IgA	€ 9.91	---	(1998)
Cost of endoscopy with biopsy	€ 386	---	
Cost of major complication	€ 6200	---	(Mein and Ladabaum, 2004)
Costs of IBS per patient per year	€ 275	Lognormal (5.61,0.03)	(Goettsch et al., 2004)
Cost of gluten free diet per patient per year	€ 1093	Lognormal (7,0.17)	(2008b)

Taking the sensitivity and specificity of the tTG test into consideration and using a prevalence of 6.15% for CD in the population screened, the chances of a true or false positive result for tTG will be 8.2 % in our base case analysis. These positive cases will undergo endoscopy in which 29% will turn out to be false positives. The remaining 71% are correctly diagnosed as having CD and will receive gluten free diet as a treatment.

Health benefits

Quality adjusted life years (QALYs) are the main outcome measurements for health benefits. As shown in table 1, the utility of a person with IBS symptoms whose test results were negative is 0.675, which is the quality of life for IBS. IBS quality of life is assumed to remain the same for the rest of patient's life (Akehurst et al., 2002). In case of undiagnosed CD, the patient will have a quality of life weight as low as 0.56 for the first four years (Gray and Papanicolas, 2010). As mentioned before, a correct diagnosis is assumed to occur after this delay period, resulting in improved quality of life from then on. Whenever a patient is diagnosed with CD and is adhering to a gluten free diet, quality of life is assumed to increase to 0.84 (Gray and Papanicolas, 2010) and stay the same for the remaining life expectancy. Therefore, the utility gain associated with a correct diagnosis and treatment is assumed to be 0.28, which is quite considerable.

Patients with undiagnosed CD have an increased mortality rate and hence, a lower life expectancy. The Standard Mortality Rate (SMR) for untreated CD is 1.6 (Shamir et al., 2006). Since CD was assumed to be detected anyway after four years, the increased mortality risk will just affect the first four future years of the patient's life. The life expectancy of an IBS patient is assumed to be the same as that of the general population of the same age. Applying a discount rate of 1.5% (2006) and using the Dutch life tables (2003), we found a discounted remaining life expectancy of 33.23 and 33.20 years for the general population and undetected CD patients at the time of screening respectively. Therefore, the survival gain from immediate correct diagnosis rather than after a four years delay will be 0.03 years.

Costs

As listed in table 1, relevant testing and treatment costs were taken into account, and these were estimated bottom-up. Cost of serological tests, endoscopy, complications, IBS care, and gluten free diet were distinguished.

The tests used in the screening strategy are conducted by various labs within the Netherlands, using different prices. Central tariffs exist, which act as a reference value. We used prices as listed on the website of Sanquin Blood Supply Foundation (1998), €39 for the tTG and an added €10 for the IgA test. Considering different tariffs for test prices, we used a range of €20-€55 for tTG test in our PSA.

Costs of endoscopy were based on estimates provided by the experts that were involved in the guideline and confirmed by the DRG tariff which was € 385, 80 in 2009.

Costs of complication were based on Mein and Ladabaum (Mein and Ladabaum, 2004). Costs of GP visits were assumed to be comparable for IBS and CD patients, that is, both types of patients by assumption visited their primary care physician with the same frequency. The specific medication and hospital costs of IBS patients were extracted from a Dutch case-control study (Goettsch et al., 2004). The costs of a gluten free diet were based on data from the Dutch CD guideline and the website of Dutch Coeliac Disease organization (2008b). Estimated costs varied, which is logic, given the variability of people's food patterns. We found a lower bound of € 253 and an upper bound of €1417 for dietary costs in 2009. The mean annual cost of diet used in base case analysis was € 1093 (2008b).

All costs are presented for price level 2009 in table 1. Price indices were used to update costs to this price level if necessary.

Sensitivity analysis

One way sensitivity analyses were performed for some important parameters and model choices. First, we examined the scenario of screening patients with all types of IBS, instead of targeted screening to see how excluding patients with constipation only has affected the results. Then, the health benefits related to a proper diagnosis of CD were investigated by estimating the cost-effectiveness of

the screening strategy assuming that diagnosis and treatment of CD had a) only a utility gain, or b) only a survival gain. Another scenario tested in sensitivity analysis was using a couple of blood tests instead of one in the screening process. We added Endomysial Antibodies (EMA) test in combination with tTG in the base case analysis. The result of the combination was assumed to be positive if either one of them showed a positive result. Therefore, such analysis would indicate if more precise but more expensive testing at the beginning of the screening strategy would lead to better cost effectiveness results.

The effect of assuming equal discount rates for health and money was also investigated in one way sensitivity analysis. We ran the model for the case in which, according to the old Dutch guidelines, both health gains and cost are discounted to 4%.

Costs of endoscopy and of complications related to endoscopy were hard to estimate. Therefore, we varied them using a broad range of 50% to 150% of baseline estimates to check the sensitivity of the outcomes for these input parameters.

Prevalence of CD in patients with IBS symptoms was changed to 3%, which is the prevalence used in the cost effectiveness study by Mein and Ladabaum (Mein and Ladabaum, 2004) based on data from two British studies (Sanders et al., 2001; Sanders et al., 2003), rather than our 4.7% value which was based on Dutch data. Furthermore results were investigated using a lower limit of 0.024 for the utility gain, based on the value used by Mein and Ladabaum(Mein and Ladabaum, 2004).

The delay period before a correct diagnosis of CD without screening was increased to 8 years, also to lifelong to check how the assumption of a correct diagnosis of CD after only four years in the absence of screening affects the cost effectiveness. We also included lower delay times to examine the cost-effectiveness of the screening in a situation in which patients are routinely tested for CD. However, such situation is usually rare and four years was considered a conservative estimate.

The prevalence of constipation in patients with IBS symptoms was changed by ±10% compared to the baseline estimate of 25%, which was based on a Dutch study (van der Veek et al., 2007) and consultation of experts.

Furthermore, a probabilistic sensitivity analysis was conducted to examine the overall effect of parameter uncertainty. Monte Carlo simulation was run using 10,000 replications, to examine the effects of parameter uncertainty and to evaluate

the robustness of the results. As shown in table 1, we have assigned probability distributions to all uncertain parameters of the model. The distribution used for each parameter was selected on the basis of suggestions in the literature (Briggs et al., 2006). The distribution parameters were calibrated to fit the 95% confidence intervals or standard deviations given in the reference sources. The number of Monte Carlo runs was validated by checking whether or not results changed after increasing the number of replications to 15,000.

Value of Information Analysis

In addition to the cost effectiveness analysis, the Expected Value of Perfect Information (EVPI) was calculated as the difference between the expected net monetary benefits of the perfect and current information. EVPI determines the value of conducting additional research and informs decision makers about the value of acquiring more precise estimates of input parameters used in a cost-effectiveness analysis (Claxton et al., 2001). Population EVPI was computed using the estimation of the number of cases eligible for screening in the Netherlands. Giving an upper bound for the value of future research, population EVPI helps the decision maker to decide whether current information is sufficient for the final policy decisions and whether further research may be required. Population EVPI was obtained from the results of Monte Carlo simulations for a willingness to pay range of 0 to 30,000 €/QALY.

Budget impact

The budget impact of implementing the screening process was calculated for a 10-years time horizon. To estimate consequences of the implementation at a population level, two different screening scenarios were investigated:

1. Catch up Scenario: Screen all prevalent IBS cases with IBS-D/mix type symptoms at the beginning and then continue with screening of new incident cases in the following years.
2. Gradual Scenario: Start screening new incident cases and continue this annually.

For obtaining the budget impact of implementing the screening process, we assumed that patients with a General Practitioner (GP) diagnosis of IBS-D/mix will be screened for CD. The reason for only including patients with a GP diagnosis of IBS is that almost everybody in the Netherlands is registered with a GP. Therefore, the GP practice is the most common location for diagnosis of IBS, and hence for CD screening in those with IBS-D/mix type symptoms. A broader screening strategy would require population based screening for IBS first, to decide on subsequent eligibility for CD testing. That would imply a considerable budget would be needed for IBS screening in the first place. This would render any screening strategy unlikely to ever become cost-effective. Hence, GP based CD testing in GP diagnosed IBS cases seems to be the most logical way of organizing a CD screening strategy in IBS.

Considering the prevalence of 1.05% for diagnosed IBS in the Netherlands (van der Linden et al., 2004), and the adult population of 10,073,028 aged 20-65 years in 2009 (2003), the target population for a catch up scenario was almost 106,000 people. Using IBS incidence of 0.56% (van der Linden et al., 2004), an annual incidence of about 56,000 cases was used for the gradual scenario.

RESULTS

Base case

Starting with a cohort of 100,000 IBS patients and excluding patients with constipation type IBS (IBS-C), 75,000 patients would be eligible to enter the screening process. Performing tTG+IgA at the beginning would determine approximately 6,251 patients who either have low IgA levels or normal IgA levels with positive tTG results. Those patients must then enter the endoscopy with biopsy phase. Performing confirmatory endoscopy with biopsy would diagnose a total of 4,380 actual cases of CD. Approximately 12 patients may experience complications due to the endoscopy process.

After conducting screening in patients with IBS-D or Mix type symptoms and treatment of the diagnosed CD patients, the average discounted quality adjusted life expectancy for each patient will be 22.8 QALYs. In comparison with no screening, this is an increase of about 0.067 QALYs for an incremental cost of €418 per patient

on average. This results in an incremental cost effectiveness ratio of about 6,200 €/QALY compared to a policy of no screening.

Sensitivity analysis

Results of one-way sensitivity analysis are shown in table 2.

If all diagnosed IBS patients are screened instead of just IBS-D/mix, an additional 25,000 patients would have to be tested, resulting in 1720 positive tTG results. This would then result in 1,115 more correctly diagnosed CD cases, but also 605 more false positives. Screening process would need an additional €300 per QALY gained for each patient compared to the base case targeted screening. The incremental cost-effectiveness ratio of adding CD screening in IBS-C would be 27,300 € /QALY, which is quite high. Therefore, our exclusion of constipation type IBS seems to be worthwhile.

In case that only survival gains are taken into account, the cost-effectiveness ratio will increase considerably. Ignoring survival gains, or assuming a utility gain of 0.024 (Mein and Ladabaum, 2004) rather than 0.28 from correct diagnosis also results in less favorable outcomes. However, the screening strategy still seemed cost effective in these cases, with incremental ratios below 20,000 €/QALY.

Double testing scenario which includes a second blood test (EMA) in combination with tTG resulted in a slightly higher cost-effectiveness ratio. Assuming an equal 4% discount rate for health and costs resulted in a small decrease in the ratio, making the results more favourable.

The period of delay that was applied for the correct diagnosis of CD in absence of screening also seems to have an important effect on the results: if patients with IBS-D/mix symptoms whose CD has not been diagnosed are assumed to be diagnosed earlier than a period of 4 years, the incremental cost-effectiveness ratio will increase significantly. However, it will not exceed the unofficial Dutch threshold of 20,000 €/QALY. Hence, even in a very unlikely case in which the delay in the correct diagnosis of CD is as small as 6 months the screening strategy would still be cost-effective. On the other hand, extending the time to correct diagnosis will lead to a more favorable cost-effectiveness ratio.

The remaining parameter values including changes in the prevalence of CD and prevalence of IBS-C, costs of endoscopy and complication costs had little effect on the outcomes.

Table 2 Sensitivity analysis results

			Incremental costs/ QALYs
Base case results			6,200
Screening patients with all IBS symptoms			6,500
Screening all IBS vs. screening IBS-Diarrhoea/mix			27,300
Assuming only Utility gain			6,431
Assuming only Survival gain			14,929
Double tests strategy			6,800
Variable	Base case value	Value in SA	
Discount rates (health, costs)	1.5% , 4 %	4% , 4%	5,800
Cost of Endoscopy with Biopsy	386 €	200 €	6,149
		600 €	6,358
Cost of major complication	6200 €	3100 €	6,239
		9300 €	6,254
Prevalence of CD in IBS	4.67 %	3 %	6,413
Utility gain from diagnosis and treatment of Coeliac	0.28	0.024	12,667
Time to correct diagnosis in case of no screening	4 years	6 months	19,000
		2 years	8,700
		8 years	3,167
		lifelong	768
Prevalence of constipation	25%	15%	6,466
		35%	6,077
		50%	5,940
Prevalence of low IgA	1:875	1:200	6,270

Probabilistic Sensitivity analysis

Figure 2 depicts the cost effectiveness plane for targeted screening versus care as usual (no screening), resulting from 10,000 samples in a Monte Carlo simulation. The mean incremental cost effectiveness ratio was 6,200 €/QALY. The figure

suggests rather low uncertainty in the results: all dots were located densely in the first quadrant. Also it can be seen that all dots were located below a reference line of 15,000 €/QALY, showing that the Incremental Cost Effectiveness Ration (ICER) is unlikely to surpass 15,000 €/QALY.

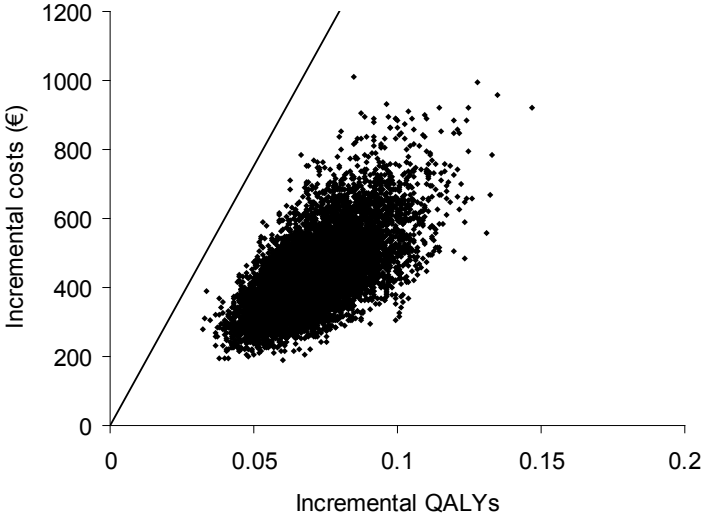


Figure 2 Results of probabilistic sensitivity analysis: Incremental cost-effectiveness ratios (ICERs) for 10,000 samples in Monte Carlo simulation. The 15,000 €/QALY line is shown as a reference

Cost Effectiveness Acceptability Curve (CEAC) is depicted in Figure 3, and again illustrates that for willingness-to-pay thresholds higher than 15,000 €/QALY, the screening strategy is almost certainly cost effective. Therefore, the decision of accepting the screening strategy seems to carry negligible risk of exceeding the threshold.

Such certainty was confirmed by the results for the Expected Value of Perfect Information. As presented in Figure 4, the maximum population EVPI peaked at about € 2.7 million at a willingness to pay threshold equal to the mean incremental cost effectiveness (6,200 €/QALY). For thresholds higher than 11,000 €/QALY, additional research would have almost no value. Thus, the base case results were shown to be sufficiently robust and there was no need to run additional analyses like partial EVPI.

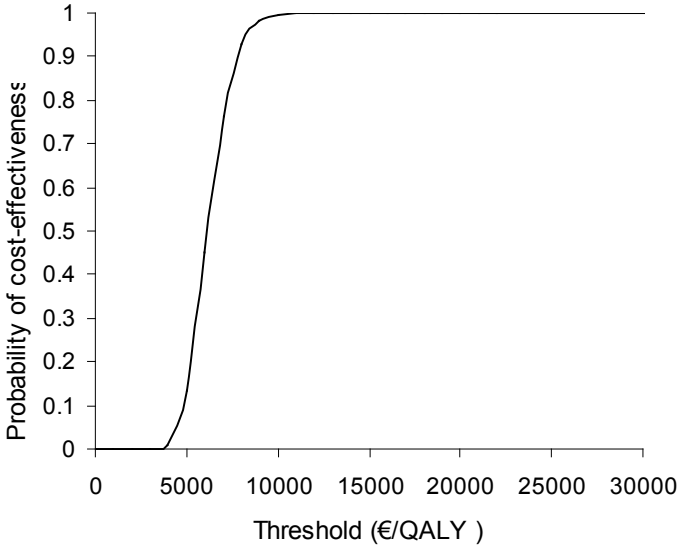


Figure 3 Cost Effectiveness Acceptability Curve for a willingness-to-pay threshold range of 0-30,000 €/QALY

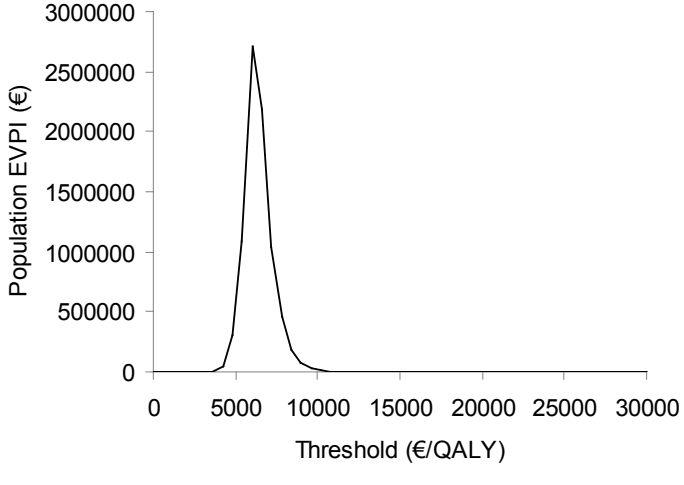


Figure 4 Population EVPI for a willingness-to-pay threshold range of 0-30,000 €/QALY

Budget impact

Screening and treatment of all patients known with IBS-D or Mix over 10 years would cost in total 25 million Euros. Costs of testing and endoscopies together would be about 6.7 million Euros. The remaining additional costs (18.3 million Euros) are incurred by treatment of CD being more expensive than treatment of IBS. In the gradual scenario, when only incident cases are considered, the total 10-years budget needed would be about 103 million Euros.

Adding these annual costs of incident cases for the next 10 years to the costs of screening current patients, the catch up scenario would cost about 128 million Euros.

DISCUSSION

The results of our study indicate that screening for CD in patients with IBS type D or mixed is a cost-effective way of improving quality of life/health for patients with IBS symptoms compared to the current situation with no structured testing strategy. The initial tTG test detects potential cases of CD with an acceptable specificity, but the number of false positives is rather high. Endoscopy with biopsy is therefore applied for a definite diagnosis of CD. Hence, costs and burden of dietary care are minimized by finding the exact group of patients with CD. However, the budget needed for conducting the screening process in The Netherlands in still high due to the high prevalence of IBS.

Three studies evaluated the cost-effectiveness of screening for CD in patients with IBS (Mein and Ladabaum (Mein and Ladabaum, 2004), NICE IBS guideline (2008a) Spiegel et al. (Spiegel et al., 2004)). In contrast to our study, the first two studies have included patients with any IBS type rather than only patients with IBS-D or IBS-Mix. The ICER of adding screening for IBS-C versus screening only mixed/D IBS would be 27,300 €/QALY. Thus, we illustrated that excluding patients with IBS-C will result in a significantly lower cost-effectiveness ratio. Since the prevalence of CD in patients with IBS-C is very low, screening them will yield little health gain. Conversely, adding these patients to the target population would unnecessarily increase the costs per QALY gained and especially the total budget impact.

The third study (Spiegel et al., 2004) evaluated the cost-effectiveness of screening for CD in patient with IBS-D. The outcomes and costs were measured differently in this study. Due to a lack of validated utility measures, the authors have used symptom improvements as a measure of health gains. Since recently new studies were published with quality of life estimates for CD and for IBS, as well as a validation of the questionnaires used in the specific patient populations, use of QALYs was now possible.

In the NICE guideline (2008a) only survival gain was taken into account. However, it has been shown that patients with diagnosed CD will have higher health utilities than undiagnosed cases (Gray and Papanicolas, 2010). Therefore, we have included the utility gains in the outcome measurements. On the other hand, Mein and Ladabaum (Mein and Ladabaum, 2004) have only taken into account the utility gain, while studies show that patients with untreated CD have a higher mortality rate (Shamir et al., 2006). Hence, such an assumption seems to be quite conservative. We have included the increase in mortality. In order to avoid overestimation of health gains, we have restricted the low utilities and survival rates of undiagnosed CD to the first four years of the analysis. Because of continued symptoms the patient is assumed to be diagnosed with CD after these four years regardless of screening.

Since incremental cost effectiveness ratios in Spiegel et al. (Spiegel et al., 2004) are stated in cost per additional symptomatic improvement and hence are inconsistent with our measurements, we cannot compare these to our results. Using the same currency and price levels in the two more consistent studies, the base case incremental cost effectiveness ratio in Mein and Ladabaum (Mein and Ladabaum, 2004) and NICE IBS guideline (2008a) were 5,800 €/QALY and 16,300 €/QALY, respectively. The main causes of the difference between cost effectiveness ratios in the two mentioned studies and our results may be the inclusion of dietary costs and differences in health benefits included. Costs of gluten free diet were considered in NICE IBS guideline (2008a) , while these costs were ignored in Mein and Ladabaum (Mein and Ladabaum, 2004). Our study used larger health benefits, since both utility and survival gains were included. This as well as a lower cost for endoscopy with biopsy in the Dutch health system may be the reason for such a difference. As shown in one way sensitivity analysis, if only survival gains are

accounted, the cost effectiveness ratio of our study will also reach values of about 14,900 €/QALY, while the exact cost figures applied were relatively unimportant for the outcomes. .

When restricting the type of health gains from a correct diagnosis and treatment of CD to only utility or only survival gains, the cost effectiveness ratios of screening was higher. Such an increase is to be expected, since ignoring some of the effects of the screening strategy will result in a less favorable cost-effectiveness ratio.

Our estimation of the utility gain from correct diagnosis and treatment of CD was different from Mein and Ladabaum's estimate. We have considered a utility gain of 0.28 QALYs based on the Gray and Papanicolas (Gray and Papanicolas, 2010) study. Mein and Ladabaum (Mein and Ladabaum, 2004) have calculated the utility gain from correct diagnosis and treatment of CD as the difference between utilities of treated CD and IBS utility. Such calculation does not seem to be quite accurate, since IBS utility is taken into account instead of untreated coeliac utility. In other words, Mein and Ladabaum (Mein and Ladabaum, 2004) have assumed that patients with IBS symptoms who are not screened for CD have the health state utility of IBS. They have ignored the probability that patients with IBS symptoms actually suffering from CD will have a lower quality of life compared to true IBS patients. We have taken into account the difference between treated and untreated CD utility, as well as the proportion of IBS patients who may have undiagnosed CD.

Time to remain undiagnosed without screening is also important. The longer the time period before a patient is detected with CD without screening, the higher the potential QALY gain for the screening strategy. Hence, the screening strategy will have higher incremental effects and a lower cost-effectiveness ratio. We used a rather conservative estimate of 4 years of delay based on the Dutch guideline comity's advice. Studies have shown that the gap between the start of symptoms and the CD diagnosis is usually more than 10 years (Fasano, 2003) and sometimes a CD patient remains undiagnosed for the rest of her life (Spiegel et al., 2004). As shown in our sensitivity analysis, longer time periods would result in more favourable, i.e. lower incremental cost-effectiveness ratios. Therefore, the cost-effectiveness ratio of our study may be considered a conservative upper bound of the possible ratio. Plausible alternative assumptions regarding the lag time before correct diagnosis

would not change the results unfavourably.

Another advantage of the current study is the elaborate inclusion and presentation of parameters uncertainty and the resulting uncertainty in the outcomes. Cost-effectiveness acceptability curves help the decision maker to decide more easily and to get an estimation of the risk of the decision when there is uncertainty in the results. Value of information analysis gives the decision maker an upper bound for the value gained by reducing the uncertainty in the input parameters.

Since there was no literature on the mean age of screening and time to remain undiagnosed in absence of screening, these parameters were valued based on expert opinions. This can be mentioned as a limitation of our study, since the assumptions are affecting our results. However, assumptions seem quite reasonable, and they prevent results from being too optimistic.

Our model is evaluating a cohort of 34 year old patients, and could be refined to evaluate a typical IBS population from 20 to 65 years old. Changing the cohort age will alter the remaining life expectancy which is multiplied by the costs and effects per year to estimate the overall cost effectiveness ratio. Since such factor equally affects both nominator and denominator of the cost-effectiveness ratio, making changes in that would not substantially alter the results.

We also ignored the costs of organizing a screening program. However, since GPs can refer for serological testing during their normal IBS consultations and then refer for endoscopy if needed, such costs are relatively low, provided the screening for CD takes place for GP diagnosed IBS only.

Endoscopy and biopsy have a very small probability of complications. If the patient survives, a short period with decreased quality of life may be assumed. Comparing to the whole life of the patient, such decrement will take place for a very short while (2-3 weeks). Since it moreover has a very low chance of occurrence, the survival difference was ignored in the analysis.

Results of our study support the notion that screening may be implemented in the Netherlands to improve the quality of life in patients labelled with IBS at relatively low costs. These results were incorporated in the new Multidisciplinary Dutch guideline for IBS.

Our study evaluates one of the possible ways of detecting CD in general population. New screening tools such as screening based on genetic rather than serologic tests could also be evaluated in future studies. For the moment the costs of these tests in comparison to their sensitivity and specificity seem to favour the serologic tests applied in the current analyses. To conclude, testing patients with diagnosed IBS-D or IBS-mix for CD is almost certainly cost-effective. Assumption regarding type and size of health gains from CD diagnosis affects the results, while time to remain undiagnosed without screening is also important. However, results were quite robust.

REFERENCES

Dutch Coeliac Disease organization. http://www.glutenvrij.nl (2010). Accessed 2 Aug 2010

Nederlandse Zorgautoriteit (Dutch Healthcare Authority). http://dbc-tarieven.nza.nl/Nzatarieven/top.do. Accessed 2 Aug 2010

Sanquin Blood Supply Foundation. http://www.sanquin.nl (1998). Accessed 6 Apr 2010

Statistics Netherlands. www.cbs.nl (2003). Accessed 14 June 2010

Richtlijnen voor farmaco-economisch onderzoek, geactualiseerde versie (Guidelines for pharmacoeconomic research, updated version). College voor zorgverzekeringen (College of Health Insurances, CVZ) (2006)

Irritable bowel syndrome in adults: Diagnosis and management of irritable bowel syndrome in primary care (Clinical practice guideline). National Institute for Health and Clinical Excellence: 2008a.

Richtlijn Coeliakie en Dermatitis Herpetiformis (Coeliac disease and dermatitis herpetiformis guideline). Nederlandse Vereniging van Maag-Darm-Leverartsen (Dutch Association of Gastroenterologists and Hepatologists): 2008b.

Akehurst RL, Brazier JE, Mathers N, O'Keefe C, Kaltenthaler E, Morgan A. Health-related quality of life and cost impact of irritable bowel syndrome in a UK primary care setting. Pharmacoeconomics 2002; 20(7): 455-462.

Berg MJt, Goettsch WG, Boom Gvd, Smout AJ, Herings RM. Quality of life of patients with irritable bowel syndrome is low compared to others with chronic diseases. European Journal of Gastroenterology & Hepatology 2006; 18(5): 475-481.

Bracco A, Jonsson B, Ricci JF, Drummond M, Nyhlin H. Economic evaluation of tegaserod vs. placebo in the treatment of patients with irritable bowel syndrome: an analysis of the TENOR study. Value in Health 2007; 10(4): 238-246.

Brazier J, Dolan P, Karampela K, Towers I. Does the whole equal the sum of the parts? Patient-assigned utility scores for IBS-related health states and profiles. Health Economics 2006; 15(6): 543-551.

Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics 2002; 21(2): 271-292.

Brazier J, Usherwood T. Deriving a preference based single utility index from the UK SF36 health survey. Journal of Clinical Epidemiology 1998; 51(11): 1115-1128.

Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. Oxford University Press: Oxford;2006.

Bushnell DM, Martin ML, Ricci JF, Bracco A. Performance of the EQ-5D in patients with irritable bowel syndrome. Value in Health 2006; 9(2): 90-97.

Claxton K, Neumann PJ, Araki S, Weinstein MC. Bayesian Value of Information Analysis: An Application to a Policy Model of Alzheimer's Disease. International Journal of Technology Assessment in Health Care 2001; 17(1): 38-55.

Fasano A, Berti, I, Gerarduzzi, T, Not, T, Colletti, RB, Drago, S, Elitsur, Y, Green, PH, Guandalini, S, Hill, ID,

Pietzak, M, Ventura, A, Thorpe, M, Kryszak, D, Fornaroli, F, Wasserman, SS, Murray, JA, Horvath, K. Prevalence of celiac disease in at-risk and not-at-risk groups in the United States: a large multicenter study. Archives of Internal Medicine 2003; 163(3): 286-292.

Goettsch WG, van den Boom G, Breekveldt-Postma NS, Smout AJPM, Herings RMC. Treatment patterns and health care costs of mebeverine-treated IBS patients: a case-control study. Pharmacology and drug safety 2004; 13: 803–810.

Gray AM, Papanicolas IN. Impact of symptoms on quality of life before and after diagnosis of coeliac disease: results from a UK population survey. BMC Health Services Research 2010; 10: 105.

Hershcovici T, Leshno M, Goldin E, Shamir R, Israeli E. Cost effectiveness of mass screening for coeliac disease is determined by time-delay to diagnosis and quality of life on a gluten-free diet. Alimentary Pharmacology & Therapeutics 2010; 31: 901-910.

Hungin APS, Whorwell PJ, Tack J, Mearin F. The prevalence, patterns and impact of irritable bowel syndrome: an international survey of 40 000 subjects. Alimentary Pharmacology & Therapeutics 2003; 17: 643–650.

Inadomi JM, Fennerty MB, Bjorkman D. Systemic review: the economic impact of irritable bowel syndrome. Alimentary Pharmacology & Therapeutics 2003; 18: 671-682.

Mein SM, Ladabaum U. Serological testing for coeliac disease in patients with symptoms of irritable bowel syndrome: a cost-effectiveness analysis. Alimentary Pharmacology & Therapeutics 2004; 19: 1199-1210.

Sanders DS, Carter JC, Hurlstone DP, Pearce A, Ward AM, McAlindon ME, Lobo AJ. Association of adult coeliac disease with irritable bowel syndrome: a case-control study in patients fulfilling ROME II criteria referred to secondary care. Lancet 2001; 358: 1504–1508.

Sanders DS, Patel D, Stephenson TJ, Ward AM, McCloskey EV, Hadjivassiliou M, Lobo AJ. A primary care cross-sectional study of undiagnosed adult coeliac disease. European Journal of Gastroenterology & Hepatology 2003; 15: 407-413.

Shahbazkhani B, Forootan M, Merat S, Akbari MR, Nasserimoghadam S, Vahedi H, Malekzadeh R. Coeliac disease presenting with symptoms of irritable bowel syndrome. Alimentary Pharmacology & Therapeutics 2003; 15: 231-235.

Shamir R, Hernell O, Leshno M. Cost-Effectiveness Analysis of Screening for Celiac Disease in the Adult Population. Medical Decision Making 2006; 26: 282–293.

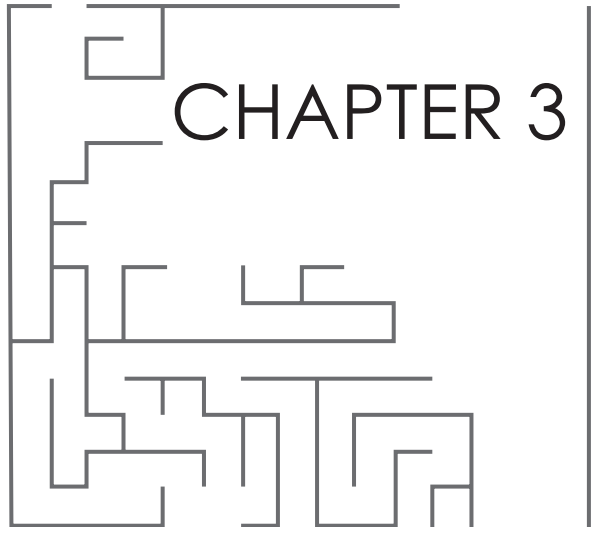
Spiegel B, Harris L, Lucak S, Mayer E, Naliboff B, Bolus R, Esrailian E, Chey WD, Lembo A, Karsan H, Tillisch K, Dulai G, Talley J, Chang L. Developing valid and reliable health utilities in irritable bowel syndrome: results from the IBS PROOF Cohort. American Journal of Gastroenterology 2009; 104(8): 1984-1991.

Spiegel BM, DeRosa VP, Gralnek IM, Wang V, Dulai GS. Testing for celiac sprue in irritable bowel syndrome with predominant diarrhea: a cost-effectiveness analysis. Gastroenterology 2004; 126(7): 1721-1732.

van der Linden M, Westert GP, De Bakker DH, Schellevis FG. Tweede Nationale studie naar ziekten en verrichtingen in de huisartspraktijk. Klachten en aandoeningen in de bevolking en in de

huisartspraktijk (The second Dutch national survey of general practice. Complaints and disorders in the population and in general practice). NIVEL:2004; Utrecht.

van der Veek PPJ, van Rood YR, Masclee AAM. Clinical trial: short- and long-term benefit of relaxation training for irritable bowel syndrome. Alimentary Pharmacology & Therapeutics 2007; 26: 943–952.



Value of information analysis from a societal perspective: a case study in prevention of major depression.

Leyla Mohseninejad

Pieter van Baal

Matthijs van den Berg

Erik Buskens

Talitha Feenstra

Value in Health. 2013 ;16(4):490-7

Abstract

Objectives: Productivity losses usually have a considerable impact on cost-effectiveness estimates while their estimated values are often relatively uncertain. Therefore, parameters related to these indirect costs play a role in setting priorities for future research from a societal perspective. Until now, however, value of information analyses has usually applied a health care perspective for economic evaluations. Hence, the effect of productivity losses has rarely been investigated in such analyses. The aim of the current study therefore was to investigate the effects of in- or excluding productivity costs in value of information analyses.

Methods: Expected Value of Information Analysis (EVPI) was performed in cost-effectiveness evaluation of prevention from both societal and healthcare perspectives, to give us the opportunity to compare different perspectives. Priorities for future research were determined by Partial EVPI. The program to prevent major depression in patients with sub-threshold depression was opportunistic screening followed by Minimal Contact Psychotherapy (MCP).

Results: The Expected Values of Perfect Information (EVPI) indicated that regardless of perspective further research is potentially worthwhile. Partial EVPI results underlined the importance of productivity losses when a societal perspective was considered. Furthermore, priority setting for future research differed according to perspective.

Conclusion: The results illustrated that advises for future research will differ for a health care versus a societal perspective and hence the value of information analysis should be adjusted to the perspective that is relevant for the decision makers involved. The outcomes underlined the need for carefully choosing the suitable perspective for the decision problem at the hand.

INTRODUCTION

Estimates of cost-effectiveness are surrounded by uncertainty. Reduction of uncertainty is usually costly. A Value of Information (VOI) analysis estimates the monetary value of investments that may be required to eliminate all or part of uncertainty in the evaluations. Such estimation can support the decision maker in deciding whether further research is warranted (Raiffa, 1968). When further research turns out to be worthwhile, more detailed value of information analysis can identify the uncertainties which should then become research priorities. Value of information for parameters estimates the expected values of information on groups of parameters and assists the decision maker to decide on those uncertainties. The concept of value of information analysis (Raiffa, 1968) was applied in many sectors (Yokota and Thompson, 2004) before it was introduced in health technology assessment by Claxton et al. (Claxton, 1999). Recently, the number of applications in health care has steadily grown. A range of studies were published after 2004 (Black et al., 2009; Bojke et al., 2008; Claxton and Sculpher, 2006; Fox et al., 2007; Hassan et al., 2010; Rojnik and Naversnik, 2008; Smits et al., 2010; Wailoo et al., 2008; Welton et al., 2008; Wilson et al., 2010).

While the societal perspective is recommended for economic evaluations in many countries (Claxton et al., 2010) the majority of the previous studies have applied a health care perspective in analyzing the value of information. Studies performed in the UK were just following national directives in adopting a health care perspective according to The guidelines manual (2009). Also among the non-UK studies, however, only a few have taken into account other than direct health care costs. Some included direct non-healthcare or some part of indirect healthcare costs (Smits et al., 2010; Spronk et al., 2008), however, they have ignored productivity losses. Galani et al. (2008) mentioned that cost estimates included indirect costs, but they did not elaborate further on the consequences of this for the interpretation of their results. Nevertheless, most guidelines that recommend using a societal perspective also suggest comparing the results from two perspectives. Such a comparison has been usually missing from the studies. The review by Yokota and Thompson (Yokota and Thompson, 2004) highlights that also in other sectors

applications have often chosen a relatively narrow perspective. Hence, to the best of our knowledge, only very few papers have included productivity losses and if they did, the implications were not thoroughly discussed. This seems an omission, since in interventions that target chronic diseases with a high prevalence among patients in their working ages, productivity costs may have a large impact on cost-effectiveness results. Furthermore, productivity costs often can only be estimated with large uncertainty. Information on individuals' working hours as well as hourly productivity may be difficult to ascertain and is not often included in most clinical trials. Therefore, looking at the impact of the choice of perspective and inclusion of productivity costs on the outcomes of a value of information analysis is worthwhile and was the aim of the current study.

A case study was chosen in the field of mental disorders. Many reports show that mental disorders lead to a reduction in employee productivity due to absenteeism or impaired functioning at work (Dewa et al., 2007). Depression is one of the major mental disorders, with a high burden of disease (Kuijshaar et al., 2005; Mykletun et al., 2009). Due to work loss, absenteeism and presentism, productivity losses resulted from depression are considerable (Smit et al., 2006). A recent study showed that productivity costs, on average, reflect more than half of the total costs for treatment of depressive disorders (Krol et al., 2011). In fact, the majority of costs of depression fall outside the health care sector, i.e., the benefits of preventing depression are not restricted to the health sector but society as a whole. Accordingly, for many health care decision makers it will be relevant to consider a societal perspective in addition to the health care perspective in evaluating cost effectiveness in depression prevention. Still, to date most economic evaluations of treatments for adults with depressive disorders have ignored productivity losses (Krol et al., 2011).

The objective of our study then was to perform a value of information analysis in cost-effectiveness evaluation of preventing major depression in patients with minor depression. We considered both a societal and a health care perspective and paid attention to the consequences of different perspectives for policy advice. The depression case serves as an illustration for many interventions with large but uncertain effects on productivity costs.

MATERIALS AND METHODS

In a recent study (VanDenBerg et al., 2011) a Markov model based on Vos et al. (Vos et al., 2005) was used to evaluate the costs and long term health benefits of screening followed by Minimal Contact Psychotherapy (MCP) for depression prevention. The model was adjusted to allow evaluation of depression prevention and was adapted to the Dutch setting. The short term outcomes of MCP were previously evaluated alongside a randomized controlled trial (Smit et al., 2006). The current chapter adds an elaborate value of information analysis, and focuses on a comparison between the values of solving uncertainties for different perspectives. The model was used to extrapolate the trial outcomes over a five-year time horizon. Five years were considered long enough to capture the full effects of the intervention and still short enough to trust the data on the population and the screening results. The discount rates used were 1.5% and 4% according to the Dutch Guidelines for pharmacoeconomic research (2006), and monetary outcomes were valued in Euros, at the 2008 price level. For clarity reasons, we explain the intervention, the model and parameter estimation sources in the following sections.

Intervention

The intervention was opportunistic screening for sub-threshold depression followed by Minimal Contact Psychotherapy (MCP). Full details about the intervention and its short term effects compared to no screening have been published before (Willemse et al., 2004). In short, opportunistic screening takes place in three steps: first, people are approached by the assistant when they are in the waiting room during a regular GP visit Those who are eligible for screening and give informed consistent (participation rate: 72.5%) are then screened for sub-threshold depression (screen positive rate: 26.6%). In a second step, screen-positive patients are approached for a further screening to check whether they meet the inclusion criteria for sub-threshold depression (participation rate: 35.7%). Those who meet all inclusion criteria receive MCP (59.5% of positive screens).

MCP consists of a self-help manual with instructions on cognitive-behavioral self-help in mood management skills. The manual contains registration exercises and homework assignments aimed at cognitive restructuring, relaxation, and

activity scheduling to increase pleasant activities.

In the control group no screening took place. People with sub-threshold depression received care as usual from their GP, i.e., they were offered treatment upon presenting themselves with symptoms.

The effects of the intervention were twofold: incidence and recurrence of major depression decreased by 6% (Willemse et al., 2004) and the total annual per capita costs decreased by 21% (Smit et al., 2006).

Patient population

The intervention targets patients with minor (sub-threshold) depression. Sub-threshold depression, which is diagnosed when a patient has 2-4 symptoms of major depression, has a lifetime prevalence of 10% (Kessler et al., 1997). People with minor depression have an increased risk of developing major depression compared to those not meeting the criteria of sub-threshold depression (Cuijpers and Smit, 2004).

Markov model

The model distinguishes three main states: sub-threshold depression, major depression (MD) and recovered from depression (no MD). Each state is divided into episodes which last for four weeks. After each cycle of four weeks, a person has the chance of moving to another state of disease, or to stay in the same state and start a new episode within that state. The Markov model is depicted in Figure 1.

The probability of developing major depression for people with sub-threshold depression (the incidence rate) has been assumed to be independent of the time that persons were in the sub-threshold state, while the probabilities of recovery from major depression and relapse into major depression by assumption decreased over the time which was spent in MD and no MD states respectively. Parameters related to costs and QALYs in the recovered states are by assumption the same as in the sub-threshold states.

Modeling and analyses were all done by means of the R software environment for statistical computing.

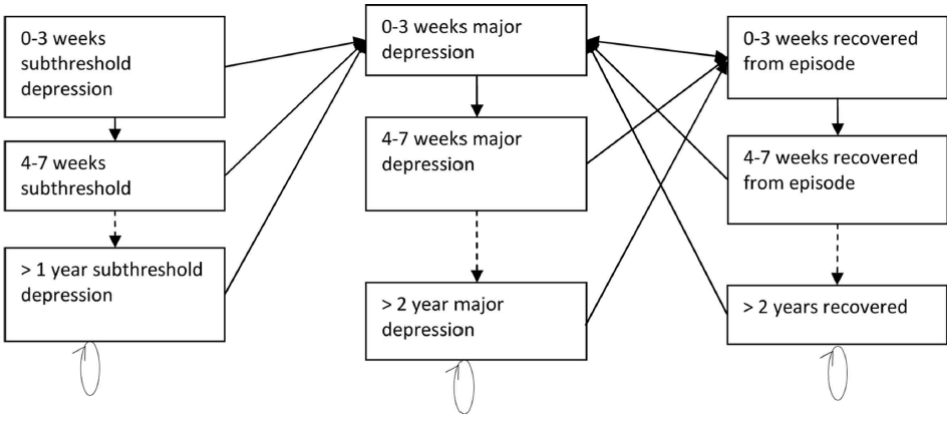


Figure 1 The Markov model

Parameter estimation

Estimates of relapse and recovery rate as a function of duration were based on the Dutch NEMESIS study (Kuijshaar et al., 2005), a large population based cohort study addressing mental disorders. The time-dependant probability curves for relapse and recovery rates can be found in van den Berg (VanDenBerg et al., 2011). Prevalence of sub-threshold depression, intervention costs, health care and societal costs for sub-threshold depression were based on trial results (Smit et al., 2006). Population parameters and incidence probabilities from sub-threshold to major depression were taken from Willemse et al. (Willemse et al., 2004). Costs and productivity losses for major depression were estimated based on a review of Dutch studies (Bosmans et al., 2007; Stant et al., 2008; VanRoijen et al., 2006). QALY estimates were based on the NEMESIS study (Kuijshaar et al., 2005). Distribution functions were estimated for all important model parameters. Parameters that were used in the probabilistic sensitivity analysis are presented in Table1.

Effects of MCP were conservatively assumed to cease one year after the intervention. That is, after one year, persons are assumed to return to the same risk of developing major depression as under care as usual if still in the sub-threshold state.

Table 1 Distributions of the parameters of intervention used in PSA (EVPPI group indicates how the parameters were grouped for the partial value of information analysis)

		No MCP		MCP		EVPPI group
		Distribution	Mean (s.e.)		Mean (s.e.)	
Screening process	Fraction of the target population that agrees to be screened [26]			Beta distribution (alfa=3826; beta=1452)	0.725 (0.006)	Population
	Fraction of screened included for diagnostic interview [26]			Beta distribution (alfa=364; beta=3463)	0.095 (0.005)	Population
	Fraction of interviewed included in intervention [26]			Beta distribution (alfa=217; beta=148)	0.595 (0.026)	Population
Sub threshold states	Incidence probabilities from sub-threshold to major depression [26]	Beta distribution (alfa=21; beta=90)	0.016 (0.003)	Beta distribution (alfa=14; beta=95)	0.011 (0.003)	Incidence rate
	Health care costs [21]	Gamma distribution (shape=15; scale=108)	1627 (419)	Gamma distribution (shape=31; scale=55)	1687 (305)	Health care costs sub-threshold
	Productivity loss [21]	Gamma distribution (shape =22; scale=300)	6481 (1393)	Gamma distribution (shape =8; scale=576)	4638 (1634)	Productivity loss
	Direct non medical costs [21]	Gamma distribution (shape =43; scale=12)	507 (77)	Gamma distribution (shape=56; scale=8)	441 (59)	----
	Quality of life [19]	Uniform distribution (0.81-1)	0.91 (0.05)	Same as No MCP		QALYs
Major depression states	Health care costs [29,30,31]	Gamma distribution (shape=15; scale=152)	2280 (589)	Same as No MCP		Health care costs major depression
	Productivity loss [29,30,31]	Gamma distribution (shape=8; scale=27)	216 (76)	Same as No MCP		Productivity loss

Expected Value of Perfect Information

The Expected Value of Perfect Information (EVPI) was calculated for a cost effectiveness range of 0 to 60,000 €/QALYs. The global EVPI was computed as the difference between the expected net benefit of perfect and current information over a sufficient number of simulations, and the net benefits of the standard therapy (No MCP) were assumed to be zero:

$$EVPI = E_{\theta} \max[0, NB(MCP, \theta)] - \max\ [E_{\theta} NB(MCP, \theta), 0].$$

Here θ represents a list of unknown parameters. Population EVPIs were then computed by multiplying global EVPIs by the relevant population sizes. These were based on the prevalence of sub-threshold depression (Smit et al., 2006).

In addition to this, parameters were grouped to find the Expected Value of Perfect Parameter Information (EVPPI) for each of these groups of parameters. EVPPI or partial EVPI is intended to inform research priorities, that is, the type of additional evidence which would be the most valuable to inform the decision. Parameters which explained the same concepts were grouped together as shown in Table 1. Like the global EVPI, the Partial EVPIs were analyzed for different cost-effectiveness thresholds to see how they varied for a range of thresholds.

The EVPPI is calculated as the difference between the expected value of a decision made with perfect information on a group of parameters and the expected value with current information on that group of parameters. It reflects the maximum value of additional information on the value of this group of parameters and may serve to help decide whether or not certain research to find better information on the parameters is worth its costs:

$$EVPPI_{\varphi} = E_{\varphi} \max\ [0, E_{\psi|\varphi} NB(MCP, \varphi, \psi)] - \max\ [E_{\theta} NB(MCP, \theta), 0].$$

Where φ is the group of parameters of interest and ψ represents the remaining uncertain parameters. To compute the partial EVPI, first the simulation must be run for parameters ψ but with a particular value of φ (an inner loop) and then a new

value of φ is sampled and the simulation is run again (an outer loop). This process is repeated until we have sampled sufficiently from the distribution of φ (Briggs et al., 2006).

Number of simulations

Careful selection of the number of simulations needed in the inner and outer loop is required to balance off computation time and precision. The numbers of sufficient inner and outer loops for the EVPPIs were computed using a three stage algorithm which estimates the bias and confidence intervals for the outcomes of EVPPI (Oakley et al., 2010). Predicted bias and the width of 95% confidence intervals for different number of inner loops (*J*) and outer loops (*K*) are presented in Table 2 for the health care perspective. The numbers presented are relative values, indexing the global EVPI (43,500,000 €) to 100. We chose J=1000 and K=100 as the sufficient numbers for our simulation from both perspectives, since the bias was reasonably low and also the width of confidence interval was estimated to be low enough at 3% of the global EVPI.

Table 2 Predicted bias and 95% CI for Monte Carlo partial EVPI estimate, EVPI indexed to 100

	J=10	J=100	J=500	J=1000	J=5000	J=10000
Bias (indep't of K)	7.62	-0.54	-0.21	-0.05	0.03	-0.09
95% CI						
K=10	303.74	56.63	24.82	17.95	11.1	5.2
K=100	35.94	10.14	4.5	2.94	1.51	0.83
K=500	7.16	2.02	0.9	0.59	0.3	0.17
K=1000	3.58	1.01	0.45	0.29	0.15	0.08
K=5000	0.72	0.2	0.1	0.06	0.03	0.02
K=10000	0.35	0.1	0.05	0.03	0.01	0.01

RESULTS

Estimates of effects and costs per intervention together with incremental effects and costs are shown in Table 3. From a health care perspective, the incremental cost effectiveness ratio of MCP compared to the standard therapy (no MCP) is

about 1100 €/QALY. From a societal perspective, the intervention is cost-saving (VanDenBerg et al., 2011).

Table 3 Estimates of effects and costs per intervention and incremental effects and costs

	QALYs *(1,000)	Health care costs *(1,000,000)	Total costs *(1,000,000)
No MCP	1158	2911	11612
MCP	1170	2924	11240
Incremental	12	13	-372

We illustrate our results showing outcomes for two different perspectives in single pictures/graphs with the cost-effectiveness threshold on the horizontal axis, allowing a comparison of the societal and health care perspective.

EVPI

Figure 2 depicts the Cost-Effectiveness Acceptability Curves (CEACs) together with results of EVPI from a societal and a health care perspective. CEACs were also presented by van den Berg (VanDenBerg et al., 2011), but are repeated here to support explanation of the EVPI.

From a health care perspective, at low cost effectiveness thresholds, prevention using MCP is not cost effective. The value of perfect information is relatively low, since the decision not to implement MCP is relatively certain to be the best decision. The EVPI rises to a maximum of 57 million Euros at a threshold value of about 1100 €/QALYs, which is equal to the mean cost-effectiveness of the intervention from a health care perspective. For larger thresholds, the probability that the MCP is cost effective increases and the EVPI decreases with the threshold rising. At a threshold surpassing the mean incremental cost-effectiveness ratio, we expect the intervention to be cost-effective and the decision is less likely to be changed by further information. With increasing cost effectiveness thresholds the value paid for each additional QALY increases. Hence, at high values of the threshold, the global EVPI rises again due to an increased investment risk.

In contrast, from a societal perspective, the intervention is cost saving on average and the probability of a correct decision is always increasing as the threshold gets higher; hence the global EVPI is always decreasing. We expect, however, that at very high thresholds (which are not shown in these graphs), the EVPI starts to rise, because high threshold values mean that a wrong decision is extremely costly.

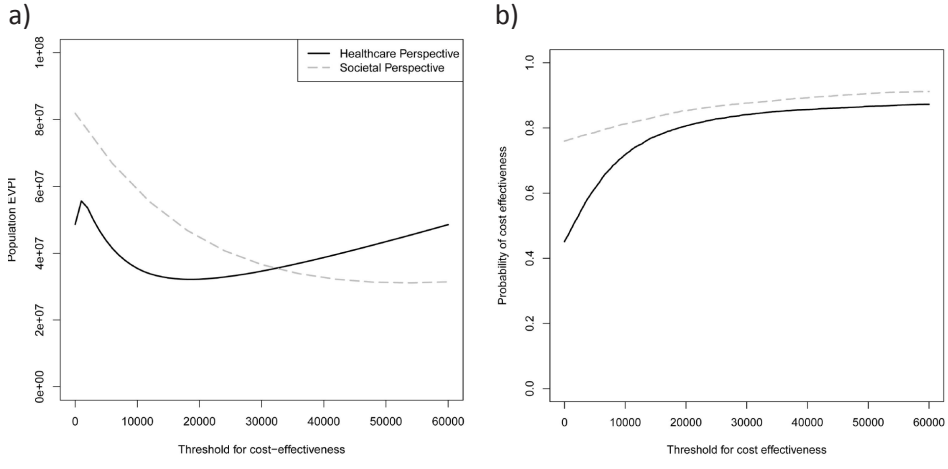


Figure 2 (a) Global expected value of perfect information (EVPI) curves; and (b) Cost effectives acceptability curves

It is apparent that the value of information is mostly higher from a societal perspective comparing to a health care perspective. Towards the right end of the willingness to pay-scale, however, the VOI from health care perspective rises above the levels of the EVPI from a societal perspective. This can be explained by having a look at the CEAC: it shows that the probability of making an incorrect decision remains higher for the health care perspective. Comparisons, however, of the value of information for different perspectives might require different willingness to pay thresholds. The amount that the decision maker is willing to pay per additional QALY would most likely change when different perspectives are considered. Therefore, to reach the best comparison, the values of the thresholds relevant for the decision makers should be known. For instance, if the threshold for evaluating MCP from a health care perspective is 20,000 €/QALY and the threshold for the same intervention from a societal perspective is 40,000 €/QALY, then the CEACs and EVPIs must be compared on two different points of 20,000 and 40,000 €/QALY

on the x-axis. The vertical lines in Figure 2 show the comparison considering these hypothetical threshold values. We will come back to this point in discussion.

Partial EVPI

The expected value of perfect information for separate parameters groups from the societal perspective is illustrated in Figure 3. Results indicated that when the societal perspective is considered, the productivity loss was the most important source of uncertainty at any threshold. The effect of productivity loss, however, was more important for low thresholds than for higher ones. The next important parameter group was health care costs (including both sub-threshold and major depression), which became more prominent at higher thresholds. The third and fourth priority would be given to parameters related to QALYs and the incidence rate. Other parameters, such as population and recovery-relapse rates were not significantly affecting the value of information, indicating that they would have a low priority in further research.

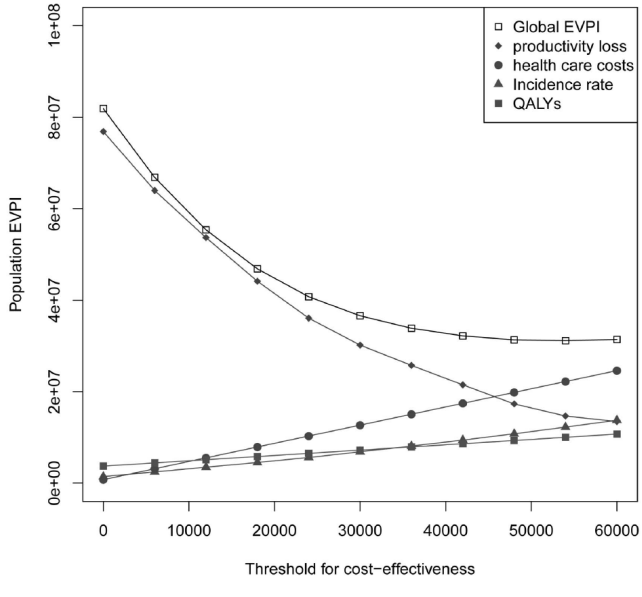


Figure 3 Partial EVPI curves considering a societal perspective.

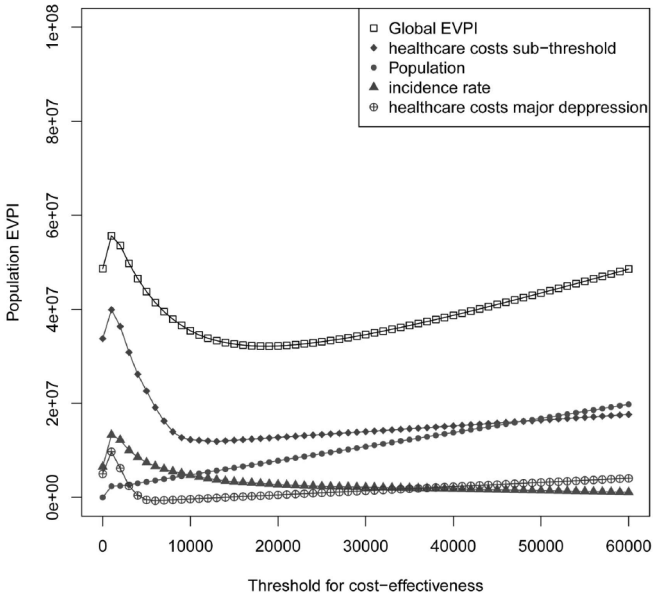


Figure 4 Partial EVPI curves considering a health-care perspective.

Results of the partial EVPI from a health care perspective are shown in Figure 4. The graph indicated that the most important sources of uncertainty were health care costs. Since health care costs had such a clear effect on the uncertainty of the problem, they have been subdivided into two groups of parameters: health care costs in sub-threshold and in major depression. At very low thresholds costs of major depression were more important than at high thresholds. The health care costs of major depression had an almost negligible expected value of information at thresholds above 5000 €/QALYs. The incidence rate also had considerable impact at low thresholds, but low impact at high thresholds. Results showed that the most important priorities in future research were those related to the health care costs of sub-threshold depression. The population which will finally receive the MCP also became an important parameter at high willingness to pay thresholds. This parameter is determined by the fraction of the target population that agrees to be screened, fraction of screened included for diagnostic interview and fraction of interviewed included in intervention. Other parameters, related to QALYs, recovery and relapse rates were of very little importance in the health care perspective.

DISCUSSION

This chapter examined how the value of information would change for different perspectives. The case study was the decision whether or not to use opportunistic screening in combination with MCP to reduce the incidence of major depression. For depression, absence from work and the associated productivity costs represent an important part of the burden of disease. For this reason, results were evaluated both from a societal perspective and from a health care perspective. We found that regardless of the perspective, parameters related to costs had the largest expected values of partial information, that is, resolving their uncertainty would be most valuable for this case study. From a societal perspective, however, productivity costs got priority, while these were by their very nature ignored from a health care perspective.

The current case study could be illustrative for many other mental disorders: often productivity costs represent a relatively large part of the disease burden, and they are also often relatively uncertain due to a lack of data. It is obvious that when the societal perspective is relevant for the decision maker, a VOI from a health care perspective may lead to erroneous priorities for further research, especially in presence of large and uncertain productivity costs.

Comparing the cost-effectiveness results from a societal perspective to a health care perspective as required by most guidelines recommending a societal perspective, indicates that including societal costs in the analyses may significantly affect the outcomes and even change the decisions. The changes occur not only regarding the acceptability of the intervention but also regarding priorities for further research and value of information. According to the cost-effectiveness acceptability curves (Fig. 2), for very low willingness-to-pay thresholds from a health care perspective rejecting the intervention seems to be the most reasonable decision. From a societal perspective, however, for low thresholds accepting the intervention has a fair chance of being cost-effective.

To have an estimation of the EVPI we look at the graphs at the unofficial Dutch thresholds for preventive interventions of 20,000 €/QALY (VanDenBerg et al., 2008). This threshold was first mentioned in a health care perspective setting (Casparie

et al., 1998). It has also been used, however, for analyses in a societal perspective (VanDenBerg et al., 2008). At the threshold of 20,000 €/QALY, the global EVPIs had a value of about 42 and 32 million Euros from societal and healthcare perspectives respectively, both indicating that it would be worthwhile to gather more information.

Hence, using an invalid perspective could lead to unrealistic importance attached to additional research, depending on the actual threshold value. As mentioned in the results, however, considering the same threshold for both perspectives is not very practical. In real world decisions, threshold values change based on the perspective chosen. From a societal perspective threshold values should reflect the consumption value of health, while from a health care perspective they would reflect the marginal value of health provided for by a publicly financed health care system, which are not necessarily the same. Opinions differ, however, on this issue (Claxton et al., 2010; Jönsson, 2009). Meanwhile, it is not clear how large the difference between the two thresholds should be. In applications, similar threshold values are sometimes mentioned in studies applying a societal perspective as well as studies using a health care perspective. For instance, as mentioned before, the unofficial Dutch threshold of 20,000 €/QALY has been used in both health care and societal perspectives, indicating the difficulties in understanding the relation between perspective and threshold in many decision contexts.

We choose to illustrate our results in graphs with a single threshold on the x-axis. If the actual thresholds were larger from a societal perspective (following the reasoning in Claxton et al. (Claxton et al., 2010)) then a figure which compares both perspective should use two different scales on the x-axis, effectively shifting the societal perspective graph to the left. This implies that the differences between the societal and health care perspective in the value of the global EVPI decrease. It is obvious that the only way to reach a precise comparison between the VOI from the two perspectives is to know the exact willingness-to-pay threshold considering each perspective.

Using a health care perspective, value of information analysis informs decision makers about allocating funding for actual interventions and research that basically can be considered as originating from the same health care budget. All costs and savings hence refer to the same budget and decision maker, even if in reality

earmarking and separate budgets will be present. Using a societal perspective, however, this may no longer hold. For instance, savings in productivity costs will accrue to employers, not to health care decision makers. In recent years this has led to discussions on how the costs of interventions and research must be sponsored when a societal perspective is considered. Debates are mostly focused on public health interventions, in which the impacts are often wide-ranging. Costs and benefits associated with an intervention aiming at public health, like the depression prevention case, will fall on many sectors within the society (Weatherly et al., 2009). Some authors just assume that when a societal perspective is taken, the society pays for health care interventions through a single payer system and also pays for research projects for reducing uncertainties through government or private donation based agencies (Willan and Pinto, 2005). Weatherly et al. (Weatherly et al., 2009) reviewed a number of approaches that have been suggested to account for the impact of interventions across different sectors. For instance, Claxton et al. (Claxton et al., 2007) introduce a multi-sectoral societal decision-making approach to evaluate costs and benefits which fall on different sectors of the economy. Smith et al. (Smith et al., 2005) also demonstrate the value of using a macroeconomic approach to modelling a major health problem, using the context of antimicrobial resistance and applying general equilibrium analysis. Following Willan and Pinto (Willan and Pinto, 2005), it seems valid to argue that a societal perspective implies that resources can –in principle- be transferred from one part of the economy to the other and the extended Pareto criterion may be applied to decide whether an intervention or additional research is worthwhile. Hence, in presence of uncertainty, value of information analysis will inform whether additional research is potentially worthwhile from a societal point of view to support better future decision making, independent of who is going to pay or gain from this research.

To conclude, our results underlined the need for carefully choosing the relevant perspective for the decision problem at the hand, also in value of information analyses, in order to avoid erroneous choice of research priorities.

REFERENCES

Richtlijnen voor farmaco-economisch onderzoek, geactualiseerde versie. Available from: www.cvz.nl. [Accessed March 11, 2010].

The guidelines manual. National Institute for Health and Clinical Excellence 2009; London.

Black C, Clar C, Henderson R, MacEachern C, McNamee P, Quayyum Z, Royle P, Thomas S. The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation. *Health Technology Assessment* 2009; 13(52): 1-148.

Bojke L, Claxton K, Sculpher M, Palmer S. Identifying Research Priorities: The Value of Information Associated with Repeat Screening for Age-Related Macular Degeneration. *Medical Decision Making* 2008; 28: 33-43.

Bosmans JE, Brook OH, VanHout HP, DeBruijne MC, Nieuwenhuyse H, Bouter LM, Stalman WAB, VanTulder MW. Cost effectiveness of a pharmacy-based coaching programme to improve adherence to antidepressants. *Pharmacoeconomics* 2007; 25(1): 25-37.

Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation*. Oxford University Press: Oxford;2006.

Casparie AF, VanHout BA, Simoons ML. Guidelines and costs. *Nederlands Tijdschrift voor Geneeskunde* 1998; 142: 2075-2077.

Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 1999; 8(3): 341-364.

Claxton K, Sculpher MJ. Using Value of Information Analysis to Prioritise Health Research: Some Lessons from Recent UK Experience. *Pharmacoeconomics* 2006; 24(11): 1055-1068.

Claxton K, Sculpher MJ, Culyer A. Mark versus Luke? Appropriate methods for the evaluation of public health interventions. CHE Research Paper 31. Centre for Health Economics:2007; York.

Claxton K, Walker S, Palmer S, Sculpher MJ. Appropriate perspectives for health care decisions. *CHE Research Paper 54* 2010:

Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatrica Scandinavica* 2004; 109: 325-331.

Dewa CS, McDaid D, Ettner SL. An international perspective on worker mental health problems: who bears the burden and how are costs addressed? *Canadian Journal of Psychiatry* 2007; 52(6): 346-356.

Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, Taylor RS. The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model. *Health Technology Assessment* 2007; 11(47): iii-iv, ix-248.

Galani C, Al M, Schneider H, Rutten FF. Uncertainty in decision-making: value of additional information in the cost-effectiveness of lifestyle intervention in overweight and obese people. *Value in Health* 2008; 11(3): 424-434.

Hassan C, Pickhardt PJ, Giulio ED, Hunink MGM, Zullo A, Nardelli BB. Value-of-Information Analysis to Guide Future Research in the Management of the Colorectal Malignant Polyp. *Diseases of the Colon & Rectum* 2010; 53(2): 135-142.

Jönsson B. Ten arguments for a societal perspective in the economic evaluation of medical innovations. *European Journal of Health Economics* 2009; 10: 357–359.

Kessler RC, Zhao S, Blazer DG, Swartz M. Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. *Journal of Affective Disorders* 1997; 45(1-2): 19-30.

Krol M, Papenburg J, Koopmanschap M, Brouwer W. Do Productivity Costs Matter? The Impact of Including Productivity Costs on the Incremental Costs of Interventions Targeted at Depressive Disorders. *Pharmacoeconomics* 2011; 29(7): 601-619.

Kuijshaar ME, Barendregt J, Vos T, DeGraaf R, Spijker J, Andrews G. Lifetime prevalence estimates of major depression: An indirect estimation method and a quantification of recall bias. *European Journal of Epidemiology* 2005; 20: 103-111.

Mykletun A, Bjerkeset O, Overland S, Prince M, Dewey M, Stewart R. Levels of anxiety and depression as predictors of mortality: the HUNT study. *British Journal of Psychiatry* 2009; 195(2): 118-125.

Oakley JE, Brennan A, Tappenden P, Chilcott J. Simulation Sample Sizes for Monte Carlo Partial EVPI Calculations. *Journal of Health Economics* 2010; 29(3): 468-477

Raiffa H. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Random House: New York;1968.

Rojnik K, Naversnik K. Gaussian process metamodeling in Bayesian value of information analysis: a case of the complex health economic model for breast cancer screening. *Value in Health* 2008; 11(2): 240-250.

Smit F, Willemse G, Koopmanschap M, Onrust S, Cuijpers P, Beekman A. Cost-effectiveness of preventing depression in primary care patients Randomised trial. *British Journal of Psychiatry* 2006; 188: 330-336.

Smith RD, Yago M, Millar M, Coast J. Assessing the macroeconomic impact of a healthcare problem: The application of computable general equilibrium analysis to antimicrobial resistance. *Journal of Health Economics* 2005; 24: 1055–1075.

Smits M, Dippel DWJ, Nederkoorn PJ, Dekker HM, Vos PE, Kool DR, VanRijssel DA, Hofman PAM, Twijnstra A, Tanghe HLJ, Hunink MGM. Minor Head Injury: CT-based Strategies for Management—A Cost-effectiveness Analysis. *Radiology* 2010; 254(2): 532-540.

Spronk S, Bosch JL, Ryjewski C, Rosenblum J, Kaandorp GC, White JV, Hunink MG. Cost-effectiveness of new cardiac and vascular rehabilitation strategies for patients with coronary artery disease. *PLoS ONE* 2008; 3(12): e3883.

Stant AD, TenVergert EM, DenBoer PC, Wiersma D, Groen H. Cost-effectiveness of cognitive self-therapy in patients with depression and anxiety disorders. *Acta Psychiatrica Scandinavica* 2008; 117: 57-66.

VanDenBerg M, DeWit GA, Vijgen SM, Busch MC, Schuit AJ. Cost-effectiveness of prevention:

opportunities for public health policy in the Netherlands. *Nederlands Tijdschrift voor Geneeskunde* 2008: 152:1329-1334.

VanDenBerg M, Smit F, Vos T, VanBaal PHM. Cost-Effectiveness of Opportunistic Screening and Minimal Contact Psychotherapy to Prevent Depression in Primary Care Patients. *PLoS ONE* 2011; 6(8): e22884. doi:22810.21371.

VanRoijen LH, VanStraten A, Al M, Rutten F, Donker M. Cost-utility of brief psychological treatment for depression and anxiety. *British Journal of Psychiatry* 2006; 188: 323-329.

Vos T, Corry J, Haby MM, Carter R, Andrews G. Cost-effectiveness of cognitive-behavioural therapy and drug interventions for major depression. *Australian and New Zealand Journal of Psychiatry* 2005; 39(8): 683-692.

Wailoo AJ, Sutton AJ, Cooper NJ, Turner DA, Abrams KR, Brennan A, Nicholson KG. Cost-effectiveness and value of information analyses of neuraminidase inhibitors for the treatment of influenza. *Value in Health* 2008; 11(2): 160-171.

Weatherly H, Drummond M, Claxton K, Cookson R, Ferguson B, Godfrey C, Rice N, Sculpher M, Sowden A. Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. *Health Policy* 2009; 93: 85-92.

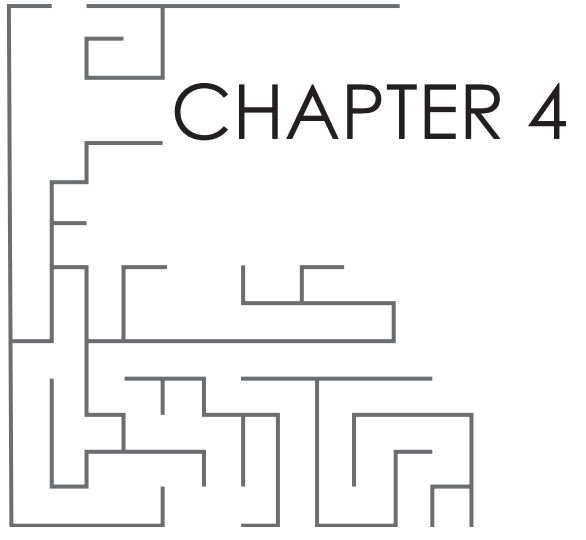
Welton NJ, Ades AE, Caldwell DM, Peters TJ. Research prioritization based on expected value of partial perfect information: a case-study on interventions to increase uptake of breast cancer screening. *Journals of the Royal Statistical Society* 2008; 171, part 4: 807–841.

Willan AR, Pinto EM. The value of information and optimal clinical trial design. *Statistics in Medicine* 2005; 24: 1791–1806.

Willemse G, Smit F, Cuijpers P, Tiemens BG. Minimal-contact psychotherapy for sub-threshold depression in primary care Randomised trial. *British Journal of Psychiatry* 2004; 185: 416-421.

Wilson E, Gurusamy K, Gluud C, Davidson BR. Cost–utility and value-of-information analysis of early versus delayed laparoscopic cholecystectomy for acute cholecystitis. *British Journal of Surgery* 2010; 97: 210-219.

Yokota F, Thompson KM. Value of Information Literature Analysis: A Review of Applications in Health Risk Management. *Medical Decision Making* 2004; 24: 287-298.



Optimal timing of the "adoption with research" period for conditional reimbursement of drugs.

Leyla Mohseninejad

Talitha Feenstra

Kuno Huisman

Maarten Postma

Cornelis Boersma

Erik Buskens

Submitted.

Abstract

Access under the condition of evidence development plays an increasing role in drug reimbursement decisions. The goal is to ensure quick market access and deployment of innovations while simultaneously assuring collection of further evidence collected to support informed decision making. In the Dutch setting for inpatient drugs, conditional reimbursement implies that drugs are provisionally reimbursed, while prospective observational research is undertaken during a predetermined period. This period has been fixed at four years, followed by a re-evaluation and definite reimbursement decision. The period of four years is arbitrary and the question is what period is optimal and whether it should be case specific. The current study presents a framework for establishing a suitable time point for making a definite decision. This framework takes a decision makers' approach. Deriving the expected value of the additional information, and trading it off against the costs of delaying the definite decision, the expected net gains from delay can be found and optimized. The optimal period for adoption under research is shown to be variable depending on several drug- and indication specific parameters.

INTRODUCTION

In recent years, new challenges have risen due to the advent of new innovative but expensive drugs. While these drugs are potentially effective for some individuals, their overall cost-effectiveness may be questionable given limited healthcare resources (Nikolentzos et al., 2008). Therefore, it is essential to carefully select evaluation criteria for these drugs. Where health technology assessment (HTA) is part of decision-making systems, cost-effectiveness studies usually are a onetime event, used when making the reimbursement decision for a certain drug (Martikainen and Rajaniemi, 2002). Though these analyses involve sensitivity analysis (univariate and/or probabilistic) and scenario analysis to give insight in the uncertainty level, decisions often ignore that part of the uncertainty around estimates of costs and effects will dissolve over time. That is, the relevant decision is not just whether to reimburse or not, but also when and under what conditions to start reimbursement of a new drug.

Ideas for regulation have been introduced that deal with this aspect of timing in drug reimbursement decisions in different jurisdictions. For instance in the UK, a system has been proposed distinguishing four categories of guidance for new medications. Based on the estimated cost-effectiveness, degree of uncertainty, and costs of further research, a new medication may be rejected, approved, approved with research (AWR), or recommended only in research (OIR) (Claxton et al., 2011).

Alternatively, but very similarly, Eckermann and Willan (Eckermann and Willan, 2007) distinguished three decision options: 1) Adopt and no trial (AN): To adopt the new medication with no further research; 2) Delay and trial (DT): To delay the decision in order to search for more evidence using a clinical trial. 3) Adopt and trial (AT): Adopt the new medication and search for new evidence, using a clinical trial with certain size and fixed final time.

The “Approval with Research”/“Adopt and Trial” category is comparable to what is called “conditional reimbursement” in the Netherlands. Conditional reimbursement has actually been put to practice. It was implemented as of 2007 for expensive intramural drugs, and then was extended in 2012 to all drugs claiming added value compared to existing treatments. Drugs are temporarily reimbursed

while more evidence on the added value and cost-effectiveness is gathered. After a period of four years (Boer, 2012) the new information is added to the available evidence and a definite reimbursement decision is to be made. Other countries have also introduced variants of access with evidence schemes (Carlson et al., 2010). In most European countries including the Netherlands, the reimbursement decision is made separate from the market approval, with different criteria in different countries (Martikainen and Rajaniemi, 2002). So far coverage and reimbursement of new drugs have been linked to the collection of additional evidence in many countries (Carlson et al., 2010; Mohr and Tunis, 2010; Towse and Garrison, 2010; Trueman et al., 2010), and several studies have reported on combinations of (economic) evaluation of drugs and more flexible ways of decision making (Chen and Willan, 2013; Eckermann and Willan, 2008a; Eckermann and Willan, 2007, 2008b; Willan, 2008; Willan and Kowgier, 2008; Willan and Pinto, 2006).

The period of four years for making the definite decision is arbitrarily chosen in the Netherlands, raising the question of whether a model may be developed that allows for flexible timing of the moment of re-evaluation. Such a model would have to balance several aspects. Waiting for new information may add value by reducing uncertainty. On the other hand, a longer period of conditional reimbursement will be associated with increased research costs, as well as opportunity losses resulting from more patients using a sub-optimal medication.

In the Dutch setting, data collection during the four year period is usually performed by means of observational data, for instance a patient registry. This has two reasons. First, additional trials are hard to organize in the same jurisdiction once a medication has been –even conditionally- admitted. Second, reimbursement authorities usually want additional evidence on local effectiveness and costs in actual practice, which trials are less suited to produce.

Registries also have the advantages of reflecting daily practice quite closely and of not requiring patient to agree to randomization for their treatment. Observational designs, however, have drawbacks. The most important is of course the lack of randomisation, which introduces the possibility of biased results. Especially estimating effectiveness can be complicated by selection and other biases.

These risks may be managed to some degree by good design (Dugas et al.,

2008) and proper analysis of data, e.g. by using propensity scores (Indurkhya et al., 2006). This, however, is not the focus of the current manuscript. Since the purpose of the current manuscript is to improve actual conditional reimbursement schemes in some other aspects, we will ignore any possible bias in the registry data in our first sections, while admitting such biases will affect how models perform. We turn back to the issue in the discussion.

In the current study, we develop a model which allows for finding the optimal data observation period during a conditional reimbursement process. We divide the time of the conditional reimbursement to different stages. We take a healthcare perspective, in which healthcare costs and registry costs are both covered by national authorities. We extend and modify the approach set out by Eckermann and Willan (Eckermann and Willan, 2007) to the relevant case where new evidence is gathered over time in a patient registry, while numbers of patients on each of the relevant drugs are driven by actual practice and disease characteristics. Furthermore, we deliberately account for dependency of the observations over time. Different solution strategies are considered, depending on the required timing of the decision and the follow-up period, either before the registry starts or later. Using simulation then enables us to establish the optimal solution and avoid the complexity and data requirements of a full analytic solution that would hinder actual application.

Based on the simulations, we find an optimal time point for making the final decision. This point is optimal, given the knowledge available at that point in time, i.e., waiting for further information would not be worthwhile anymore. That is because the expected losses of further delay of a decision are more than the expected gains from ascertaining more information.

In the remainder of the chapter, first a background section relates our model to the existing literature on sequential sampling, real options analysis, and multistage trials. Then in section 3, the theoretical model is set out. Section 4 presents a simple hypothetical example to clarify how the model works and to test its robustness to changes in input parameters. Finally, the discussion turns back to the issues concerning observational research mentioned above and discusses strengths and limitations of the proposed approach.

BACKGROUND

Several authors have addressed the issue of timing when deciding on new research, but this was almost exclusively in an experimental setting (trials) with little attention to everyday policy decisions and use of observational data. The current section briefly summarizes what has been published and explains what has to be added in order for the methods to be applicable in a setting of access with evidence development like the Dutch version of conditional reimbursement.

Traditional power analyses is concerned with optimal sample size (n) (Lachin, 1981), using fixed criteria on type I and type II errors. Recently, several authors used decision theory to approach the problem of optimal sample size by balancing the gains and losses from more information gathering. Willan and Pinto (Willan and Pinto, 2006) balance trial costs, opportunity losses and gains from a better informed decision to determine the optimal sample size of a single stage trial using a Value of Information (VOI) approach. Subsequent studies (Eckermann and Willan, 2008a; Eckermann and Willan, 2007, , 2008b; Willan, 2008; Willan and Pinto, 2006) have also used VOI analysis to provide models which identify optimal strategies and design for a single stage trial, using different perspectives. While they include the time horizon as a variable in their models, they do not optimize with respect to time and do not discuss the timing aspect of actual decision making.

Research on multi-stage design extends this in a way that allows for including the time aspect more explicitly in the problem. Berry and Ho (Berry and Ho, 1998) used a Bayesian decision-theoretic approach from an industry perspective and propose a procedure to stop a trial at any stage once the new treatment shows negative efficacy. Taking a societal perspective, Willan and Kowgier (Willan and Kowgier, 2008) have extended their single stage trial model (Willan and Pinto, 2006) to include multi-stage adaptive designs, while another study by the same group of authors used an industry perspective (Chen and Willan, 2013). The latter two studies (Chen and Willan, 2013; Willan and Kowgier, 2008) are relevant and related. By optimizing the number of stages along with the sample size, their model concerns the timing aspect. A major disadvantage of their approach however is its complexity. While a full analytic approach is proposed in the theoretical model,

actual application is restricted to cases with only two stages, since the complexity of the solutions increases geometrically with the number of stages. As a solution, they suggest finding simpler near-optimal solutions numerically when the numbers of stages increase, but this is left for further research.

The current study concentrates on the length of the observation period as the variable of interest, taking the number of patients as given. The idea is that this reduces complexity and will result in a tractable model that can be applied in practice.

While the studies discussed so far took a decision theoretical approach, sequential analysis has been proposed much earlier (Wald, 1945) using a traditional statistics approach. Sequential analysis assumes that a certain statistical test is performed to compare the two groups of the patients in the trial in each stage. If the null hypothesis is rejected, the trial is terminated, otherwise it continues for one more stage (Wald, 1945). This implies that criteria based on given sizes of type I and type II errors determine the number of stages.

From the field of economics comes another relevant method, the real options approach, introduced by Myers (Myers, 1984) and explained extensively by Dixit and Pindyck (Dixit and Pindyck, 1994). This approach explicitly evaluates the value of delay (rather than taking a decision right away) using similar underlying valuation of financial options (which also involves flexibility, since the holder may choose whether or not to use the option). Compared to traditional cost-benefit analysis, a real options analysis hence adds the choice of time for investment to the choice whether or not to invest.

Applying the approach to reimbursement of drugs, delay and trial (DT) can be viewed as a call option for the decision maker (payer) from the producer (applicant). Costs of doing a clinical trial is the price of that call option and the decision maker could decide whether it is worthwhile to buy that option, or to make the definite decision about the investment right away (Towse and Garrison, 2010). When the period for re-evaluation is considered flexible as we propose, the similarity is still present, yet in American options (flexible execution time) rather than European options.

In recent years, a number of studies have applied ROA to the healthcare field (Attema et al., 2010; Favato et al., 2013; Forster and Pertile, 2012; Grutters et al.,

2011; Palmer and Smith, 2000; Pertile et al., 2009) with different aims. Some of these studies analyze different investment strategies (Attema et al., 2010; Grutters et al., 2011; Pertile et al., 2009) some concern clinical decision making (Favato et al., 2013), and some incorporate ROA in decision rules in health technology assessment (Forster and Pertile, 2012; Palmer and Smith, 2000). The intention has been to use the ROA to support the healthcare decision by taking into account the different management options (Attema et al., 2010; Favato et al., 2013), the uncertainty and irreversibility associated with a technology (Forster and Pertile, 2012; Grutters et al., 2011; Palmer and Smith, 2000; Pertile et al., 2009), costs and benefits of additional research and the dynamic nature of the decision process (Forster and Pertile, 2012), the risk of investing in a suboptimal therapy (Grutters et al., 2011), or the risk of withholding patients the optimal treatment (Attema et al., 2010; Grutters et al., 2011). However, the optimal time for making a decision has not been established in any of the studies. Pertile et al. (Pertile et al., 2009) have used ROA to compare the value of immediate investment with that of postponing the investment for discrete points in time. Importantly, they were unable to find the optimal timing strategy.

A problem with applying ROA to reimbursement decisions is that the assumptions underlying the approach include full market equilibrium and interdependence over time, which will rarely be satisfied in a health care setting.

Being both decision oriented, the value of information approach and the real options approach are closely related. It has been demonstrated that (Eckermann and Willan, 2008b) the option value of delaying decisions to allow collection of further information can be estimated as the expected value of sample information (EVSI).

The decision theory approach using value of information could be considered most relevant in a health technology assessment setting. The VOI framework evaluates the additional information generated by further research, which is consistent with the objectives and constrained resources. Allowing a comparison of the potential benefits of further research with its costs, VOI provides a coherent framework for the use of health care technologies (Claxton and Sculpher, 2006). It does not require the market equilibrium assumptions and independency over time like the real options framework and has successfully been applied in the field

of health care decision making (Ades et al., 2004; Yokota and Thompson, 2004). While the differences are subtle, they concern, apart from terminology, important assumptions about the timing of information and decisions.

MODEL

To find the suitable time point for reconsidering the conditional reimbursement decision, time is divided into several stages of fixed duration. The duration of each stage may change for different drugs and conditions.

The end point of each stage is a potential point for re-evaluation. Therefore, at the end of each stage three paths are possible: continuation of the conditional reimbursement, permanent adoption of the new medication, or permanent banishment of the new drug from reimbursement. For reasons of practical feasibility, a maximum time T is set to the period of conditional reimbursement.

Assume that the new medication *A* on introduction is adopted by a fixed proportion of *r* current patients, while the rest continue using the standard of care medication *B*. Call *r* the adoption fraction. An assumption made for tractability in the model is that upon reaching the definite decision all patients change to the optimal drug, that is *r(t)=1* or *r(t)=0*, for i≥*i* (i*≤T)* where *i** is the optimal stage of making the definite decision.

INB distribution updates

For any sample, e.g. for the patients in the registry, the Incremental Net Benefits (INB) of the new drug (A) versus the old standard of care (B) can be estimated as $\lambda \times (E_A - E_B) - (C_A - C_B)$, where λ represents the willingness-to-pay threshold, E_A and E_B the mean effects of the new and old medication, and C_A and C_B are mean costs of the new and old medication, respectively. Let b_i be the estimate of *INB* at stage *i*. b_i is the parameter of interest in this model, which may be updated using real or simulated registry data after each stage.

Being a sample statistic, b_i is approximately normally distributed:

$$b_i \sim N(\hat{b}_i, \upsilon_i) \qquad i = 1,2,...,T \tag{1}$$

In the above distribution, $\hat{b}_i$ and υ_i are to be estimated with respect to the information ascertained up to the end of stage *i*. The comparison between b_i and b_{i-1} shows how the new information gained during stage *i* can help to solve part of the uncertainties in the decision and decrease the risk of making a wrong decision. We consider two approaches towards the update process, which are based on when the decision on the optimal re-evaluation time is going to be made.

Ex-ante approach

In this approach, the optimal time is to be decided upon before starting the observation period (hence at *i=0*). Therefore, the parameters of the distribution in (1) are to be estimated based on data available at that time. To find the updated INB distribution after each stage, the registry data must be simulated for all stages to evaluate how the outcomes of interest change stage by stage.

That is, histories of health benefits and costs are to be simulated for all the patients at the start of the registry and the new incident cases that enter the registry in each stage. If the population in the registry area is P_r, and *l* is the prevalence of the condition that would indicate administering drug A or B, then we will have:

$$n_{A0} = r \times P_r \times l \quad (2)$$

$$n_{B0} = (1-r) \times P_r \times l \quad (3)$$

Where n_{A0} is the number of patients who are receiving *A* and n_{B0} is the number of patients who are receiving *B* at the start of the registry (*i=0*) in the registry area. In each stage, the registry will be extended with incident cases. If we call n_{Ai} and n_{Bi} the numbers of patients whose information is available in the registry at the end of stage *i* receiving drugs *A* and *B* respectively, we have:

$$n_{Ai} = n_{A(i-1)} + (t \times r \times k \times P_r) \qquad i=1,2,\ldots,T \quad (4)$$

$$n_{Bi} = n_{B(i-1)} + (t \times (1-r) \times k \times P_r) \qquad i=1,2,\ldots,T \quad (5)$$

Where *t* is the duration of each stage and *k* is the incidence rate of the disease.

Estimating the numbers of current and new patients and using distributions for health benefits and costs, incremental net benefits can be calculated for each potential decision point. A detailed example for the ex-ante approach to find the updates in INB and solve the decision problem will be given in the example section.

Wait-and-see approach

In this approach, the point in time to make the definite decision is established as new information is arriving. Therefore, the registry population parameters as well as values for b_i can be directly observed using the registry information up to the end of i*th* stage. Such an update would still be uncertain due to the fact that the observed data sample is just one observation from a range of possible samples, but it is more precise than the ex-ante approach.

Gains of waiting as expected value of sample information

After updating b_i at the end of each stage, two outcomes are possible: $\hat{b}_i$ >0 and $\hat{b}_i$ ≤0. If $\hat{b}_i$ >0, then the decision would be between permanently adopting *A,* or continuing the registry for one more stage. Therefore in this case, the expected opportunity loss would be the loss in permanently adopting *A* from stage *i* onwards and is calculated as follows (Eckermann and Willan, 2007):

$L(b) = 0 \quad \text{if } b_i \geq 0$ (6)

$L(b) = -b_i \quad \text{if } b_i < 0$ (7)

On the other hand, when $\hat{b}_i$ ≤0, the decision is between permanently accepting *B* or continuing the registry for one more stage. Hence in this case the opportunity loss of adopting *B* is calculated as follows:

$L(b) = 0 \quad \text{if } b_i \leq 0$ (8)

$L(b) = b_i \quad \text{if } b_i > 0$ (9)

An illustration of opportunity loss of both cases is depicted in figure1.

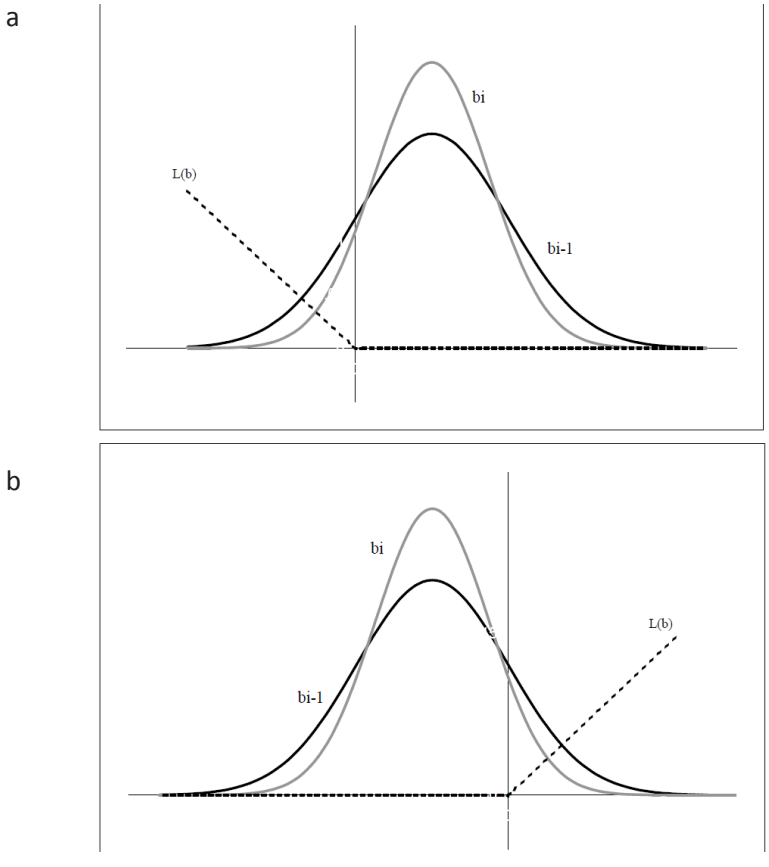


Figure 1 Updated distribution from stage i-1 to stage i and the expected opportunity loss when a) mean INB is positive at stage *i* and b) mean INB is negative at stage *i* (adapted from Eckermann and Willan 2007)

Now if we denote the total population of the country, excluding the population in the registry area by P_c, the total number of patients that can potentially benefit from a well informed decision after stage *i* can be calculated as:`

$$N_i = N_{i-1} + (t \times k \times P_c) - (t \times r \times s_A \times N_{i-1} + t \times (1-r) \times s_B \times N_{i-1}) \\ - (t \times r \times m_A \times N_{i-1} + t \times (1-r) \times m_B \times N_{i-1}) \quad (10)$$

Where m_A and m_B are the mortality rates of patients using drug A and B and s_A and s_B are the success rates of drug A and B, respectively.

At the end of each stage *i*, the Expected value of Sample Information (EVSI) is

computed as the expected reduction in the opportunity loss from end of stage *i-1* to end of stage *i* using the registry results and indicates the value gained by acquiring further information using the registry data in stage *i*. If $\hat{b}_i$ is positive, the gain of waiting one more stage versus adopting *A* is:

$$EVSI_{Ai} = N_i\left[\int_{-\infty}^{0} -b\{f_{i-1}(b)\}db - \int_{-\infty}^{0} -b\{f_i(b)\}db\right] \quad (11)$$

Where f_i is the probability density function of b_i after stage *i*.

For a negative $\hat{b}_i$, the gain of waiting for one more stage versus permanently adopting *B* would be calculated:

$$EVSI_{Bi} = N_i\left[\int_{0}^{\infty} b\{f_{i-1}(b)\}db - \int_{0}^{\infty} b\{f_i(b)\}db\right] \quad (12)$$

Costs of waiting

To compute the total costs of continuing the conditional reimbursement for one more stage when *A* is permanently adopted in the end, we include two types of costs:

1. Opportunity losses of using drug *B* in n_{Bi} patients in the registry (continuing the registry one more period), which would be calculated as $n_{Bi} \times \hat{b}_i$

2. The variable cost of being on the registry for each patient in stage *i* (C_{vi}) which is calculated as $(n_{Ai} + n_{Bi}) \times C_{vi}$

Total costs in stage *i* when $\hat{b}_i$ >0 would then be computed as:

$$TC_{Ai} = n_{Bi} \times \hat{b}_i + (n_{Ai} + n_{Bi}) \times C_{vi} \quad (13)$$

In the same way, total costs in stage *i* when $\hat{b}_i \leq 0$ would follow:

$$TC_{Bi} = n_{Ai} \times -\hat{b}_i + (n_{Ai} + n_{Bi}) \times C_{vi} \quad (14)$$

Note that the opportunity losses of using the non-optimal drug in the patients out of the registry have already been included as a part of EVSI calculations before.

Optimal decision point

The Expected Net Gain (ENG) of continuing registration rather than making a definite decision of the population in stage *i* would be:

$$ENG_i = EVSI_{Ai} - TC_{Ai} \quad \text{if } \hat{b}_i > 0 \tag{15}$$

$$ENG_i = EVSI_{Bi} - TC_{Bi} \quad \text{if } \hat{b}_i \leq 0 \tag{16}$$

Ex-ante approach

The optimal point for making the definite decision in the ex-ante approach for updating INB is obtained when the cumulative expected net gain approaches its maximum. That is, the decision maker is facing the optimization problem of finding *i* in *{0,1,2,…T}* so as to maximize $\sum_i ENG_i$. At the maximum the expected value of the additional information from continuing conditional reimbursement just balances the losses of the delaying the definite decision. A longer period of conditional reimbursement would reduce the cumulative gains.

Wait-and-see approach

Finding the optimal decision point in a wait-and-see approach requires setting some criteria on the ENG. When a registry starts, the gain in information in the first stages might be very small or even negative because it takes time to observe the full effects of the drugs. However, the registry might also continue to give poor results in the later stages. Hence, the decision maker must have a criterion to stop a non-efficient registry. Setting such criteria would need to consider the treatment time of the patient: long treatment times will require extended follow-up periods before getting some results. An analytical solution may be obtained using a Real Options Approach (Dixit and Pindyck, 1994). This, however, requires among others to assume a distribution on the INB over time. Usually this is taken to be a variant of the Brownian motion (random walk) to result in a tractable solution (Palmer and Smith, 2000). Lacking such an analytical solution, more simple rules of thumb could be applied, e.g. to stop a registry when it starts resulting in negative ENGs after having shown positive ENGs for some time.

EXAMPLE FOR EX-ANTE CASE

In this section we present the results of the model when a hypothetical example is considered. We solve the problem ex-ante, which means we will find the optimal time at the beginning of the data gathering period using simulation.

The example uses survival time as the measure of effect. Of course other effect measures could also be used. However, survival is very relevant for the setting of expensive intramural drugs, since these are often intended for treatment of severe conditions with low remaining life expectancy and have survival as an important outcome of interest. This can be readily improved by incorporating other outcome measures, depending on the case at hand.

The model as explained below applies equations (1) to (16) and has been programmed in MATLAB 7.12.0. We used 100 runs of the model to reach a robust optimal point. Larger numbers of runs did not have considerable effect on the results.

An independent graphical user interface for this model has been developed using MATLAB compiler runtime (Appendix 1)

Simulation of the registry data

Assuming the population parameters as shown in table 1 and using the formulas 2-5, we calculate the expected number of the patients at the beginning of the registry period as well as the number of incident cases per stage.

Table 1 Values for population input parameters of the example

Parameter	Basic value
Total population	16,000,000
Population of the registry area	2,000,000
Prevalence of the disease	0.05 %
Incidence proportion	0.02 %
Duration of each stage	6 months
Number of stages	16
Proportion of the patients using new drug	50%

Survival times are estimated through simulations, using the appropriate distributions. Characteristics of these distributions could be based on previous trials and other available information for each drug. For our example, we deduced the parameter values of two Weibull distributions (table 2).

The entry time for all prevalent cases at the beginning of the registry is 0. We assume random enrolment in the registry during the survival time of the prevalent patient. Therefore, assuming S_j to be the simulated survival time for patient number *j* who is already in the disease state at the start point of the registry, the disease time passed for this patient before the registry start can be simulated by:

$$p_i \sim Uniform(0, S_j)$$

And the death date can be calculated as $d_j = S_j - p_j$.

The entry time for the new incident cases is assumed to have a uniform distribution over the duration of each stage. Therefore, the entry point of an incident patient who enters the registry during stage *i* is simulated by:

$$e_i \sim Uniform(\ i-1) \times t, i \times t)$$

And the death date can be computed as $d_j = S_j + e_j$.

A mean and standard deviation for costs per time unit for each medication can be set by using drug costs from tariffs, previous data or expert opinions. Assuming some estimate of mean costs and the variability of them per drug, we assign a gamma distribution to costs per day for each simulated patient. Base case assumptions on cost parameters for each drug are also listed in table 2.

Table 2 Parameters used in registry data simulation

		Mean (SD)	Distribution
Drug A (new)	Survivals	1330 (570)	Weibull(1500,2.5)
	Costs per day (€)*	3.0(1.0)	Gamma(9.0,0.3)
Drug B (standard)	Survivals	1160(420)	Weibull(1300,3.0)
	Costs per day (€)	2.5(0.7)	Gamma(12.7,0.2)

Observed survival time in each stage

Assume that patient j, with the simulated survival time of S_j days follows the path shown in figure 2. As it is clear in the figure, when looking at the end of stage 1, the patient is censored and the observed survival for this patient is equal to t_1. At the end of stage 2, the patient is again censored, with the observed survival t_2. Finally, at the end of stage 3, the death is observed and the survival time for this patient is available. From stage 3 onwards, the record for patient *j* is complete and recorded in the database.

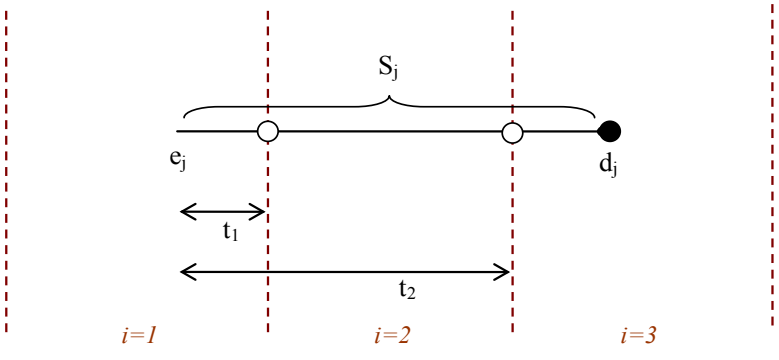


Figure 2 A hypothetical patients' disease time within different stages.

In general, assuming OS_{ji} to be the observed survival for patient *j* at the end of stage *i* with length *t* we have:

$$OS_{ij} = S_j \qquad if \quad d_j \le i \times t$$

$$OS_{ij} = i \times t - e_j \qquad if \quad d_j > i \times t$$

If the patient is already in the disease state at the beginning of the registry, observed survival time is calculated as:

$$OS_{ij} = S_j \qquad if \quad d_j \le i \times t$$

$$OS_{ij} = (i \times t) - e_j + p_j \qquad if \quad d_j > i \times t$$

Observed costs in each stage

For calculating costs in each stage, we assume that treatment costs take place during the whole life of the patient.

$$OC_{ij} = C_j \times S_j \qquad if \quad d_j \le i \times t$$

$$OC_{ij} = C_j \times (i \times t - e_j) \qquad if \quad d_j > i \times t$$

Where OC_{ij} is the observed treatment costs for patient *j* at the end of stage *i*, and C_j is the simulated costs per day of the patient *j*.

INB updates

Having the mean and standard deviation of survivals and costs as well as the number of patients in each stage *i*, the incremental net benefits can be estimated:

$$INB_i = \lambda \times \left[Mean(survival_i(A)) - Mean(survival_i(B))\right] - \left[Mean(\cos ts_i(A)) - Mean(\cos ts_i(B))\right]$$

And its variance:

$$\mathrm{var}(INB_i) = \lambda^2 \times \left[SD(survival_i(A))^2 / n_{Ai} + SD(survival_i(B))^2 / n_{Bi}\right] + \left[SD(\cos ts_i(A))^2 / n_{Ai} + SD(\cos ts_i(B))^2 / n_{Bi}\right]$$

In the base case analysis, we assume the willingness-to-pay (λ) to be 50,000 €/ life year gained and we alter this in the sensitivity analysis.

We assume that the initial distribution of *INB*, shown by INB_0, has the following rather uncertain distribution: $INB_0 \sim N(0,10000)$

Expected net gains and the optimal time

Having all simulated updates in the INB distribution for different stages and assuming registry costs of €200 per patient per year, we use the formulas 6-16 to find the expected net gains (ENG) of waiting for each stage. The optimal decision point is the point in which the cumulative ENG reaches its maximum. Figure 3 shows the average cumulative ENGs for 16 stages of 6 months each over 100 runs of the model.

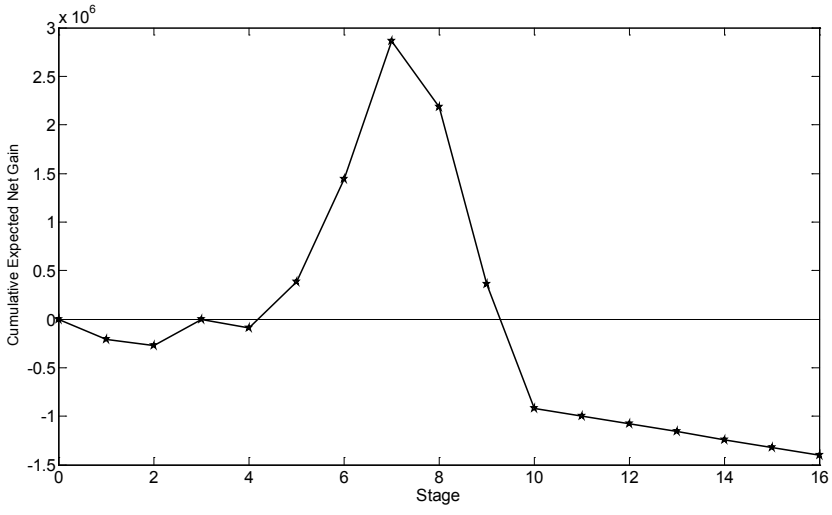


Figure 3 Average Cumulative Expected Net Gain (ENG) curve over 100 runs of the model, for a willingness-to-pay threshold of 50,000 €/life year gained and the base case parameters

As shown in figure 3, the cumulative ENG is negative in the beginning. That is mainly because in the first stages, the numbers of observed events is insufficient to get reliable survival estimates. Therefore, in the first few stages, small gains are reached from waiting for information and the costs of the registry are dominating the gains. At about stage 4, the gains of the registry start to dominate the costs and the cumulative ENG starts to rise and reaches a maximum at stage 7. Therefore, given the current parameter distribution, 3.5 years of registry would suffice to make a definite decision. After the 7th stage, the graph starts to fall. Hence, after stage 7, there would be no more gains in waiting, while in later stages more delays would cause relatively big losses.

Sensitivity analysis

We tested the sensitivity of the model to different parameters to investigate how the results can be affected by changing the model assumptions and inputs. First, we examined the effect of the assumptions about registry data. The base case analysis implies an overall INB with mean €32000 and standard deviation of €3000. Since the aim of the registry is to solve uncertainties in the estimate of INB, the standard

deviation of the data will affect the optimal time. Changing the parameters of table 2, we altered the registry INB standard deviation to the scenarios reported in table 3. A summary of the optimal points for different scenarios is shown in table 3.

Table 3 Sensitivity analysis on parameters of registry data used in data simulation

Scenario	Standard deviation of INB in the registry data	Optimal decision stage
Base case	3,000	7
Large registry data variety	10,000	9
Very large registry data variety	40,000	12

As expected, when the standard deviation of the registry data increases, the optimal time will be reached later. That is because the large variations make the information update process slower, which means that longer observation time would be needed.

We also examined the effect of important population parameters. Results of the sensitivity analysis on parameters of the population are reported in table 4.

Table 4 shows that when fewer patients are included in the registry (either because of smaller registry areas or lower prevalence and incidence proportions), a definite decision must be postponed.

Table 4 Sensitivity analysis on parameters of the population

Parameter	Base case value	Value in SA	Optimal decision stage
Population of the registry area	2,000,000	500,000	9
		15,000,000	6
Prevalence (incidence proportions) of the disease	0.05 % (0.02 %)	0.005% (0.002%)	9
		0.5% (0.2%)	6

We also examined the effect of the adoption proportion (*r*) on the optimal time. For the current settings, changing the adoption fraction did not affect the optimal point to stop the registry and make the decision. That is most probably because the mean survival time of the patients using A was not much larger than patients using

B. Therefore, the opportunity losses do not change much when more patients are using B. For different cases, for instance when one drug is much better than the other in terms of net benefits, low proportions using the better drug would lead to high opportunity losses. That would make the decision time shorter to avoid more losses.

The effect of choosing the willingness-to-pay (WTP) threshold was also examined. Results showed that high willingness-to-pay thresholds would increase both gains and losses of the registry, while low thresholds would decrease both, resulting in less risky decisions.

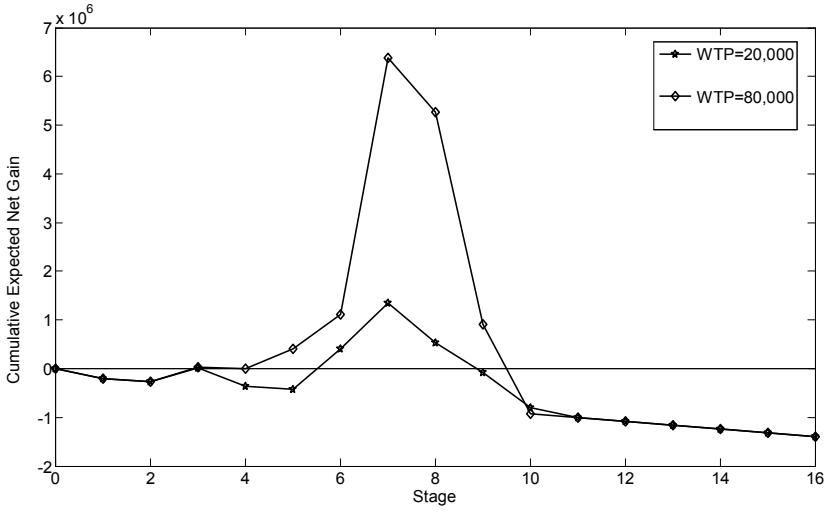


Figure 4 Average Cumulative ENGs over 100 runs of the model for two WTP threshold values.

As shown in figure 4, for a willingness-to-pay threshold as high as 80,000 €/life year gained, the curve had a higher cumulative ENG. In contrast, for a willingness-to-pay threshold as low as 20,000 €/life year gained, the cumulative ENG curve was lower but it reached the maximum at the same stage.

DISCUSSION

In this study we developed a model to describe the problem of making a definite decision on a conditionally reimbursed drug, and especially to select the optimal period of conditional reimbursement and additional evidence gathering

using observational data from a patient registry. Choosing the optimal stage for definite decisions allows the decision maker and applicant more flexibility in the conditional reimbursement setting, with additional information gathering tailored to each specific drug.

The background section described how several approaches have been put forward to deal with the issue of finding the best time of follow-up, either connected to finding the optimal sample size and using multistage trials or sequential sampling, or connected to finding the optimal time for investment decisions, using the real options approach and assuming market equilibrium.

We concluded that the value of information framework could be considered most relevant in a health technology assessment setting, and accordingly developed a model using this approach for finding the best period of conditional reimbursement and the optimal time to re-evaluate the additional data gathered during this period.

Using the proposed model, the length of the conditional reimbursement period could be set in an efficient way, tailored to the drug and disease at hand. Diseases with higher prevalence and incidence rates will require a shorter time of research. Also when a typical patient's use of a certain drug is short, less time would be needed to get to the optimal point. Another relevant variable is the size of the registry. As seen in the sensitivity analyses, when more patients are enrolled in the registry, more data will be gathered and hence the optimal time can be shortened. This indicates that larger registries could make the decision making process quicker. Of course registration costs would rise accordingly, but given that registry costs are for an important part fixed set-up costs, while per patient costs are usually more modest, efficiency gains might be possible. For instance, as a suggestion to the authorities, if regulation could impose reimbursement to be conditional on registration, the optimal time of making the definite decision could be reached faster.

Our study has several limitations that will have to be addressed before the proposed framework can be put into practice.

A first limitation may be the assumption that the additional research consists of a patient registry, and we ignore the possibility that new results from randomized controlled trials can be included in the data update process. As mentioned in the introduction however, in a setting of conditional reimbursement, recruiting patients

in the same jurisdiction to participate in a randomized trial seems infeasible. To overcome this problem, it has been suggested (Eckermann and Willan, 2009, , 2013) to use global trials where patients are recruited from a jurisdictions in which the decision has been delayed (DT). However, there is no guarantee to the reimbursement authorities that indeed such a trial would be performed and they have little instruments to impose this. Furthermore, practice variation across jurisdictions (i.e., variable inclusion criteria) may exist, resulting in similarly bias prone results. Finally, often reimbursement authorities are looking for additional information on effects and costs in actual practice rather than idealized trial circumstance. Hence, experience has shown that registries were often initiated to fulfill the requirement of additional data gathering. This same experience has also shown that many of these registries came to disappointing results at the re-evaluation time, stressing the need to carefully think over the registry design as well as the length of follow-up. Our manuscript addressed the latter point.

A second limitation of the current model is the assumption of perfect implementation. We assume that after the final decision the drug would be completely adopted or abandoned, while in practice some proportion of patients will continue using other drugs also after the final decision. Eckermann and Willan (Eckermann and Willan, 2010) relaxed this assumption for a single stage trial. For our multistage observational data model, we expect that relaxing the assumption is also possible and will not change the basic approach. However it might influence the gains and losses from delaying the decision, since it affects the opportunity losses (Eckermann and Willan, 2010). How this works out in an example is a topic for future research.

Our theoretical model is valid for any adoption proportion (*r*) and does not necessarily require it to have a fixed value over time. However, in our example we used a fixed value and analyzed different proportions in the sensitivity analysis. This could be considered a limitation, since it maybe expected that in reality the adoption proportion will vary over time, for instance showing an increasing proportion of patients using the new drug. The a-priory approach then would require to estimate this function, rather than a certain proportion, while for the wait-and-see approach, the proportion can be updated in each stage with real observations.

Actually estimating or observing these proportions might be difficult, and hence the pragmatic solution used in the example is possibly most relevant.

Varying the adoption proportion is also a way to deal with a registry which is expected to show insurmountable problems in terms of selection bias. In that case, only the data on patients that received the new drug could be used for investigating whether the original assumptions regarding the absolute effects in these patients are confirmed in practice or have to be adjusted. Then our tool may be used with r equal to 1 to find the best period for re-evaluation using the registry in this way.

In the example presented in this study we used trial based distributions to simulate future registry data. That might not be entirely appropriate, given the differences between trials and actual practice and the possible existence of learning effects. However, in absence of other data it is the best available information on what to expect from the registry.

Our model may very well result in an optimal decision time of $i^*=0$, which means the registry would not even be worthwhile to start, at least not for the purpose of informing the re-evaluation of the reimbursement decision. This might happen when overall gains are never dominating overall costs. In contrast, it might also happen that an optimal point is never reached within the time frame (T) of the problem. In other words: $i^*>T$. Such an outcome indicates that for some reason (for instance very scarce initial information on the drugs, high opportunity losses of removing a drug from the reimbursement list, low registry costs, etc), a definite reimbursement decision should be postponed to the latest possible time. Such situations require careful deliberation whether and under what conditions should a drug be allowed in reimbursement and/or a registry is the best way to inform the definite decision. It might be the case that the $i=0$ point needs to be postponed, meaning that a conditional reimbursement should not be started yet. Alternatively, different types of research maybe needed to inform the definite decision.

Our manuscript describes two methods to update the information (i.e. ex-ante or wait-and-see). Both have their limitations and strengths. In the ex-ante case, lack of real registry data would mean that the registry data must be simulated based on assumptions. These assumptions are of course uncertain and so the real optimal point might not be the same as the point found using hypothetical patients.

However, the ex-ante solution is useful in giving the decision maker an idea of the time to stop data observation in lack of any registry information. With the user interface developed along with this chapter (Appendix 1), the decision maker has the opportunity to run the model for different inputs and analyze the sensitivity to the assumptions made.

In comparison to the ex-ante approach, the wait-and-see approach uses more reliable data to obtain the optimal time for a definite decision. Such a procedure captures the full option value of delaying a definite decision (Dixit and Pindyck, 1994) and weigh it against the benefits of making a definite decision. Thus, the decision maker would not announce any specific time to the producer, but only the rules for decision making. One limitation of the wait-and-see approach is that the decision maker has to wait for data before being able to estimate a decision time. Besides, a stopping criterion for the registry must be established in the beginning of the registry period. Setting such criteria is not very straight forward in lack of an analytical solution. To improve the model and minimize the limitations of both the ex-ante and wait-and-see approaches, a combination of the two might be efficient. For instance, the decision maker can wait until some preliminary results of the registry help to find a relatively reliable estimate for the registry parameters.

Conclusion

Current adoption with research strategies need to be tailored to the specific drugs and conditions at stake. Using the model proposed in this study, the policy decision process would become more efficient. It stresses the need to carefully think about the length of the period of additional data gathering as one of the important issues to solve beforehand.

ACKNOWLEDGEMENTS

We thank Gepke Delwel and Nazanin Nooraee as well as the participants of the INFORMS 2012 and LOLA HESG 2012 meetings for their useful comments. Two anonymous referees are thanked for their suggestions that helped to seriously alter the original set up of the model.

REFERENCES

Ades AE, Lu G, Claxton K. Expected value of sample information in medical decision modelling. *Medical Decision Making* 2004; 24(2): 207-227.

Attema AE, Lugnér AK, Feenstra TL. Investment in antiviral drugs: a real options approach. *Health Economics* 2010; 19(10): 1240-1254.

Berry DA, Ho CH. One-sided sequential stopping boundaries for clinical trials: a decision-theoretic approach. *Biometrics* 1998; 44(1): 219-227.

Boer A. Procedure voorwaardelijk toelating geneeskundige zorg. College voor zorgverzekeringen (CVZ):2012.

Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy* 2010; 96: 179–190.

Chen MH, Willan AR. Determining optimal sample sizes for multistage adaptive randomized clinical trials from an industry perspective using value of information methods. *Clinical Trials* 2013; 10: 54–62.

Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, McKenna C, Soares M, Spackman E, Youn J. Uncertainty, Evidence and Irrecoverable Costs: Informing Approval, Pricing and Research Decisions for Health Technologies. CHE Research Paper 69. The University of York, Centre for Health Economics: 2011; York.

Claxton K, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics* 2006; 24(11): 1055-1068.

Dixit AK, Pindyck RS. *Investment under uncertainty*. Princeton University Press: Princeton;1994.

Dugas M, Lange M, Berdel WE, Müller-Tidow C. Workflow to improve patient recruitment for clinical trials within hospital information systems – a case-study. *Trials* 2008; 9: 2-7.

Eckermann S, Willan A. Time and Expected Value of Sample Information Wait for No Patient. *Value in Health* 2008a; 11(3): 522-526.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. *Health Economics* 2007; 16: 195–209.

Eckermann S, Willan AR. The Option Value of Delay in Health Technology Assessment. *Medical Decision Making* 2008b; 28: 300–305.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. *Health Economics* 2009; 18(2): 203–216.

Eckermann S, Willan AR. Optimal clinical trial design using value of information methods with imperfect implementation. *Health Economics* 2010; 19(5): 549-561.

Eckermann S, Willan AR. Optimal global value of information trials: better aligning manufacturer and decision maker interests and enabling feasible risk sharing. *Pharmacoeconomics* 2013; 31(5): 393-401.

Favato G, Baio G, Capone A, Marcellusi A, Saverio Mennini F. A Novel Method to Value Real Options in Health

Care: The Case of a Multicohort Human Papillomavirus Vaccination Strategy. *Clinical Therapeutics* 2013: [Epub ahead of print].

Forster M, Pertile P. Optimal decision rules for HTA under uncertainty: a wider, dynamic perspective. *Health Economics* 2012: [Epub ahead of print].

Grutters J, Abrams K, de Ruysscher D, Pijls-Johannesma M, Peters H, Beutner E, Lambin P, Joore M. When to Wait for More Evidence? Real Options Analysis in Proton Therapy. *Oncologist* 2011; 16(2): 1752-1761.

Indurkhya A, Mitra N, Schrag D. Using propensity scores to estimate the cost-effectiveness of medical therapies. *Statistics in Medicine* 2006; 26(9): 1561-1576.

Lachin JM. Introduction to Sample Size Determination and Power Analysis for Clinical Trials. *Controlled Clinical Trials* 1981; 2: 93-113.

Martikainen J, Rajaniemi S. Drug reimbursement systems in EU Member States, Iceland and Norway. Social security and health reports 54. The Social Insurance Institution:2002; Helsinki.

Mohr PE, Tunis SR. Access with Evidence Development, The US Experience. *Pharmacoeconomics* 2010; 28(2): 153-162.

Myers SC. Finance Theory and Financial Strategy. *Interfaces* 1984; 14(1): 126-137.

Nikolentzos A, Nolte E, Mays N. Paying for (expensive) drugs in the statutory system: An overview of experiences in 13 countries. London School of Hygiene & Tropical Medicine:2008; London.

Palmer S, Smith PC. Incorporating option values into the economic evaluation of health care technologies. *Journal of Health Economics* 2000; 19: 755-766.

Pertile P, Torri E, Flor L, Tardivo S. The timing of adoption of positron emission tomography: a real options approach. *Health Care Management Science* 2009; 12: 217–227.

Towse A, Garrison LP. Can't Get No Satisfaction? Will Pay for Performance Help? *Pharmacoeconomics* 2010; 28(2): 93-102.

Trueman P, Grainger DL, Downs KE. Coverage with Evidence Development: Applications and issues. *International Journal of Technology Assessment* 2010; 26(1): 79-85.

Wald A. Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics* 1945; 16(2): 117-186.

Willan A. Optimal sample size determinations from an industry perspective based on the expected value of information. *Clinical Trials* 2008; 5(6): 587–594.

Willan A, Kowgier M. Determining optimal sample sizes for multi-stage randomized clinical trials using value of information methods. *Clinical Trials* 2008; 5(4): 289–300.

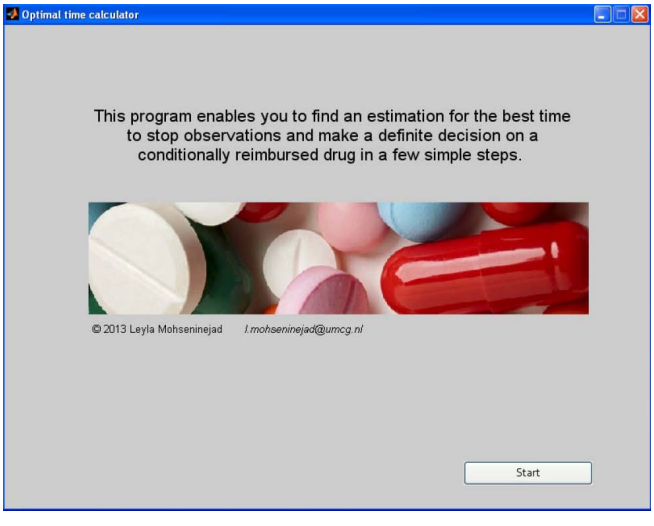
Willan A, Pinto E. The expected value of information and optimal clinical trial design. *Statistics in Medicine* 2006; 24(21): 1791–1806.

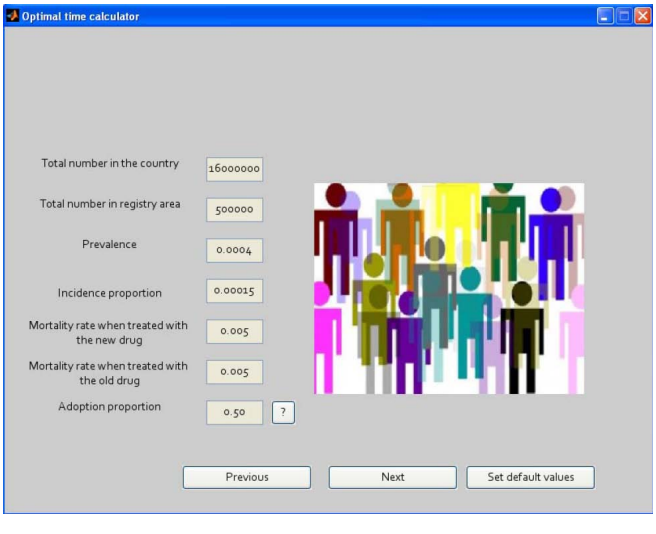
Yokota F, Thompson KM. Value of information literature analysis: a review of applications in health risk management. *Medical Decision Making* 2004; 24(3): 287-298.

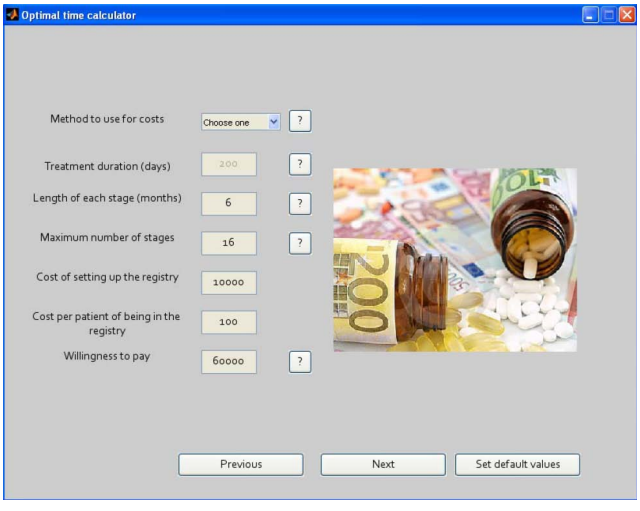
APPENDIX 1

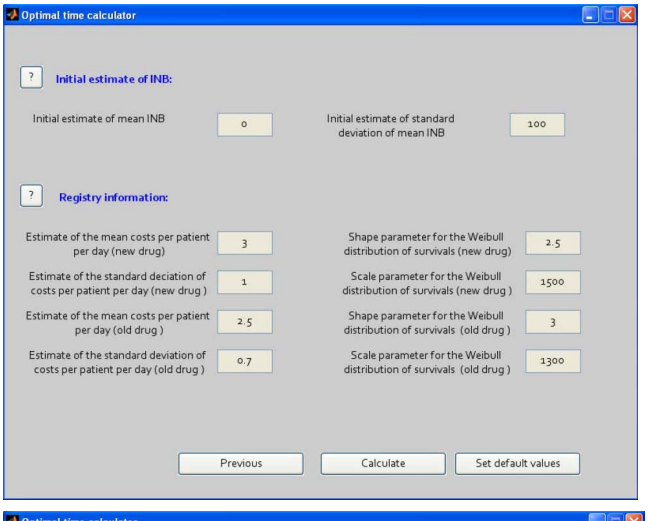
Graphical user interface

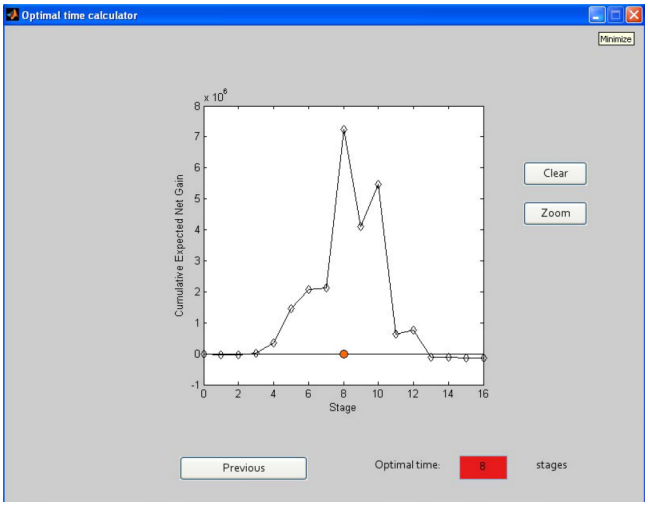
The graphical user interface developed alongside the model of this chapter allows the user to calculate an estimate for the optimal decision making time though a few simple steps. An illustration of the user interface and its consecutive steps is shown below.

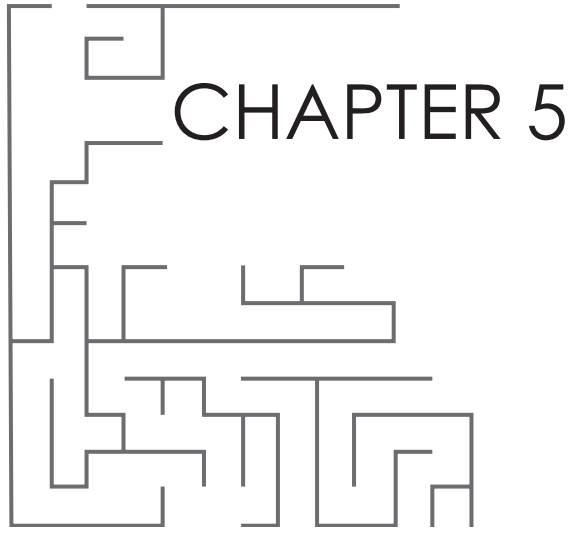





Optimal time calculator
Method to use for costs
Choose one
Treatment duration (days)
200
Length of each stage (months)
6
Maximum number of stages
16
Cost of setting up the registry
10000
Cost per patient of being in the registry
100
Willingness to pay
60000
Previous
Next
Set default values


Optimal time calculator
Initial estimate of INB:
Initial estimate of mean INB
0
Initial estimate of standard deviation of mean INB
100
Registry information:
Estimate of the mean costs per patient per day (new drug)
3
Estimate of the standard deciation of costs per patient per day (new drug)
1
Estimate of the mean costs per patient per day (old drug)
2.5
Estimate of the standard deviation of costs per patient per day (old drug)
0.7
Shape parameter for the Weibull distribution of survivals (new drug)
2.5
Scale parameter for the Weibull distribution of survivals (new drug)
1500
Shape parameter for the Weibull distribution of survivals (old drug)
3
Scale parameter for the Weibull distribution of survivals (old drug)
1300
Previous
Calculate
Set default values


Optimal time calculator
Minimize
x 10^6
Cumulative Expected Net Gain
Stage
Clear
Zoom
Previous
Optimal time:
8
stages



CHAPTER 5

When to make a definite decision regarding a conditionally reimbursed drug? The case of antifungals

Leyla Mohseninejad
Talitha Feenstra
Cornelis Boersma
Simon Daenen
Maarten Postma
Erik Buskens

Submitted.

Abstract

Background: Invasive fungal infections (IFIs) are severe complications which are difficult to diagnose and treat. Information on effectiveness, safety and cost-effectiveness of different treatment options for IFI is still scarce. Therefore, the decision of reimbursing a new antifungal drug is usually difficult to take and involves a lot of uncertainty. In The Netherlands, the Dutch reimbursement authorities used an optional procedure for expensive new inpatient drugs which involves conditional funding for a maximum period of four years along with additional data gathering. However, the 4 years length of the re-evaluation period is arbitrary.

Aim: The aim of this chapter is to find the optimum time of making a definite reimbursement decision for voriconazole as the newest drug of choice for primary treatment of invasive aspergillosis. We assume that a well-designed patient registry for outcomes research provides new observations over costs and effects of voriconazole and its comparator amphotericin B.

Methods: The optimal time is estimated a priori by distinguishing a certain number of discrete time points. Then the distribution of the incremental net benefits (INB) is derived for each time point and the expected net gains (ENG) of continuation can be computed using a value of information framework. The optimal period for adoption under research is reached when the cumulative ENG reaches its maximum.

Results: Preliminary results indicate that cumulative ENG steeply increases until about 2 years of data observation and then continues to increase slowly until it reaches an optimum at 5 years of observation. These results depend on several input parameters such as the willingness-to-pay threshold, the size of the registry and the variability of data in the registry.

Conclusion: Current conditional reimbursement procedures used in The Netherlands for expensive new inpatient drugs can be improved by using a model that enables tailoring the time to a definite decision for various drugs and conditions, using information typically available at the time of an initial decision. Gathering additional information about the cost-effectiveness of the new drug is costly, so it should be stopped when there is no more gain in waiting for new evidence. Conversely, the period of data gathering needs to be sufficiently long to obtain relevant information.

INTRODUCTION

In recent years, the discord between payers, providers and patients has intensified. Payers are responsible for ensuring prudent and principled use of scarce resources, while both providers and patients legitimately want access to technologies from which they could benefit. As a result of such a trend, new policy options for managing the uncertainty surrounding the introduction of new health technologies have emerged. For instance in the UK, the coverage options "only in research (OIR)" and "only with research (OWR)" have been proposed to guarantee further research before making a final decision (Walker et al. 2012). Such approaches typically take the form of a provisional coverage arrangement, in which the new technology is temporarily funded while evidence needed to make a definite decision is being gathered. Implementing such performance based reimbursement procedures have been an increasing trend over the last years in several countries (Carlson et al. 2010). These approaches have been referred to as 'access with evidence development' (AED) schemes (Stafinski et al. 2010). One important type of the AED scheme is called "automatic reassessment", which comprises a programmed review of a reimbursement decision following a fixed period of additional observation. This approach has become a part of the policy framework in many European countries now. However, the length of the additional research period in Europe reimbursement systems is often not specified. In some countries (Belgium, Czech Republic, Denmark, Finland, and France), a fixed time for all pharmaceuticals is set varying from 1 to 5 years. There are only a few systems in which the review period varies with the pharmaceutical (Scotland, Sweden, and UK) (Stafinski et al. 2011). In the latter systems, the time for definite decision depends upon the availability of the new evidence. However, they seem to lack a robust framework for estimating the time needed for additional research.

In the current study, we show that the length of this period can be set relatively precisely, depending on the epidemiology of the disease, costs of data gathering and other associated parameters. As a case, we investigate the conditional reimbursement of voriconazole, an antifungal drug as a treatment option for invasive fungal infections (IFIs) in The Netherlands.

IFIs are severe complications which are difficult to diagnose and treat. The incidence of infections in patients who become immuno-compromised due to chemotherapy or underlying disease is relatively high. IFIs are associated with very high rates of morbidity and mortality and therefore require adequate and timely treatment (Barnes 2008). The majority of IFIs are caused by aspergillus and candida, with aspergillus as the leading pathogen (Lehrnbecher et al. 2010).

For many years invasive aspergillus infections were treated by amphotericin B deoxycholate. However, treatment is often associated with severe toxicities that limit its use, including infusion-related reactions, nephrotoxicity, hypokalemia, and hepatotoxicity (Girois et al. 2005). Itraconazole was the first systemic azole with activity against aspergillosis, but with gastrointestinal side effects and low bioavailability of the oral compound (Heinz, Einsele 2008). Voriconazole is a relatively new broad-spectrum triazole that is active against aspergillus species (Espinel-Ingroff 2001). The voriconazole versus amphotericin B trial in 2002 (Herbrecht et al. 2002) demonstrated that initial therapy with voriconazole in patients with invasive aspergillosis would lead to improved survival and fewer severe side effects than the standard approach of initial therapy with amphotericin B. However, recent reports show the emergence of acquired resistance of aspergillus spp. to azole compounds (van der Linden et al. 2011). Also the usage of voriconazole is still limited to some extent due to its side effects, especially rising liver enzymes as well as visual disturbance and hallucination (Heinz and Einsele 2008).

Information on effectiveness, safety, and cost-effectiveness of different treatment options for IFI is still scarce, with the 2002 study (Herbrecht et al. 2002) being the only randomized clinical trial directly comparing different drugs to each other. The new antifungal drugs are usually expensive; hence it is relevant for the decision maker to know which drug offers the best value for money. However, variations in diagnostic criteria (Ascioglu et al. 2001), changes in the therapy due to adverse effects (Girois et al. 2005, Herbrecht et al. 2002, Jansen et al. 2005), and the crucial role of the environment in the epidemiology of infections (van den Bergh et al. 1999) makes the outcome measurements highly uncertain. Even adding experience from ‘real-life’ settings and observational studies might not result in sufficient evidence.

In The Netherlands, new antifungal drugs were considered to be expensive new inpatient (ENI) drugs (with costs exceeding €2.5 mln on an annual basis). In the period 2006-2011, the Dutch reimbursement authorities used a procedure for ENI drugs which involved conditional funding for a maximum period of four years. Until January 2012, the temporary funding of 80% of drug costs was based on available effectiveness data, a prognosis of cost-effectiveness and budget impact, and a plan for additional research. After the four years period, re-appraisal based on the additional information gathered during the period of conditional funding should lead to a definite decision. From 2012 on, regulations have changed. Now the expensive drugs are part of the basic health insurance, but their costs can be included as an add-on to a diagnosis-related group (DRG), allowing hospitals to ask for reimbursement of their costs in addition to the normal drug costs included in the DRG. After admission to an add-on, the reimbursement authority can assess new drugs and consider conditional reimbursement based on relevant research questions that have to be answered during the conditional period to enable re-evaluation.

The arbitrary length of the re-evaluation period was debated at the time of the ENI list. It was originally set at three years and later extended to four years. In the new regulation, this period is yet unspecified, but 4 years still seems applicable. It may be argued that the period for re-evaluation should depend on the drug and condition involved: For some drugs, it may well be that their (in)efficiency is clear at much shorter term, implying unnecessary delay of either permanent funding or exclusion from the list. For other drugs a longer period of evaluation may improve the validity of the re-appraisal.

Considering the limited evidence regarding drug effectiveness, side effects, and epidemiology of fungal infections, it seems risky to apply the simple arbitrary period of four years for gathering more evidence on the cost-effectiveness of the new antifungal drug. Therefore, it may be worthwhile to invest in determining the appropriate time for making a definite decision on an antifungal drug's reimbursement.

In this chapter, we aim to estimate the optimum time of making a definite decision on reimbursement of voriconazole as the newest drug of choice for primary treatment of invasive aspergillosis. The first decision on conditional

reimbursement of voriconazole was modelled based on data from the trial on 2002 (Herbrecht et al. 2002) as it was actually done by the Dutch Healthcare Insurance Board in 2007 (Commissie Farmaceutische Hulp (CFH) 2007). We then went on to assume that a patient registry with certain population existed and provided further information on costs and effects of voriconazole in comparison to initial therapy with amphotericin B. Such a registry has been initiated in The Netherlands. We estimate the epidemiological parameters of the condition along with distributions of costs and effects and use this to simulate the expected outcomes of the registry. We then use a value of information framework to find the gains of waiting for more information and balance these with their costs to find the optimal point of making the definite reimbursement decision on voriconazole. We also examine the effect of uncertainty in input parameters on the optimal timing.

METHODS

Patient population

We assumed that information is gathered on the prevalent cases of aspergillus spp. once the registry starts, as well as the new incident cases that are diagnosed in each time period that the registry is in place. Infection with aspergillus spp. could happen in several different categories of immunosuppressed patients. In this study, seven categories of underlying conditions were considered. These conditions, which are different types of hematologic malignancies were as follows: acute myeloid leukemia (AML), acute lymphoid leukemia (ALL), chronic myeloid leukemia (CML), chronic lymphoid leukemia (CLL), non-Hodgkin's lymphoma (NHL), Hodgkin's disease (HD), and multiple myeloma (MM) (Pagano et al. 2006).

Model

To find the gains and losses of waiting for more evidence during the conditional reimbursement period, we divided time into stages of 12 weeks length after the start point of conditional reimbursement. That is because the planned duration of fungal infection therapy is usually 12 weeks (Herbrecht et al. 2002, Denning 1996, Denning et al. 2002). Shorter stages would not be enough for observing a sufficient number

of new outcomes, while longer stages might lead to extra delays in the decision. We used the Incremental Net Benefits (INB) of voriconazole versus amphotericin B as the main variable which will define the final reimbursement decision. The INB was calculated as $\lambda \times (S_V - S_A) - (C_V - C_A)$, where λ is the willingness-to-pay threshold for each additional life week gained, S_V and S_A are survivals weeks and C_V and C_A total treatment costs of voriconazole and amphotericin B, respectively. Clearly, the final reimbursement decision would be in favour of voriconazole if INB>0 and it would be in favour of amphotericin B if INB<0.

We calculated an initial estimate for the INB based on available data at the beginning of conditional reimbursement period (explained in the next section) and then recalculated after each new stage. Data of each stage were assumed to be independent from the other stages and we used a maximum likelihood approach to find the updated INB distribution after each stage (appendix 1). Having the INB distribution in each discrete point of time, gains and losses of delaying the definite decision for one more stage were estimated. Better estimates on INBs decrease the risk in the decision; hence they result in some gain for the decision maker that may be expressed in monetary terms. Using a value of information approach, this gain was found as the Expected Value of Sample Information (EVSI). In contrast, delaying the decision causes opportunity losses, since more patients could have benefited from the optimal medication if the definite decision was made earlier. Besides, registering new patients and gathering data on them is costly. After each stage, costs of delaying the decision for one more stage were subtracted from the gains of the delay, resulting in the Expected Net Gain (ENG) of waiting for more evidence. To find the ENG of delaying the decision for each additional stage, we applied the "adopt and trial" framework suggested by Eckermann and Willan (Eckermann, Willan 2007). Their framework uses a value of sample information to estimate the gains of a trial and compares it to the costs of the trial and opportunity losses to find the ENG of the trial. We modified this framework for a registry instead of a trial, and used it repeatedly rather than once, to find the ENG of waiting in each stage. The adopted framework is described in appendix 2. The optimal point was reached when more delay in the definite decision would result in a negative ENG. This is an a priori optimization, based on the expected outcomes of the registry.

In the base case analysis we considered two scenarios: first, the case in which the prevalent patients were registered at the beginning and the incident cases were added during each stage. In this way, the number of patients using each medication in the first stage would be equal to prevalent patients who use that medication plus the new incident cases that are prescribed that medication during the first stage. After each stage, the number of patients was updated by adding the new incident cases and subtracting the dead cases. Since infected patients are usually severely ill, they were not omitted from the registry after treatment success. Even if the medication worked successfully, the patient would still be followed to observe the overall survival time as the outcome measurement. In the second scenario, prevalent patients were not included when the registry started and the information was ascertained for the incident cases only. In this way, the registry started with zero patients and new incident cases were enrolled as time passed. Like the previous scenario, after each stage new cases were added and dead cases were omitted.

The model was programmed in MATLAB and it allowed for flexible parameter inputs and comparison of results for different inputs. Since the effect of discount rates would remain limited for the time scale of the current case study, we left it out in the base case analysis. We examined a discount rate of 3% in a sensitivity analysis.

Initial Incremental Net Benefits (INB)

In order to estimate the INB at time 0, which is the time that the conditional reimbursement has started, we assigned distributions to the effect and cost estimates on which the conditional reimbursement decision was based. Health effects were defined using the 12 weeks survival proportion and mean survival, taken from the cost-effectiveness analysis (Jansen et al. 2005) based on the initial trial (Herbrecht et al. 2002) and used in the original decision on conditional reimbursement. A Weibull distribution was calibrated to fit the survival times of the two arms of the trial. Costs were also taken from the cost-effectiveness study (Jansen et al. 2005) and updated to the price level 2008. A gamma distribution was fitted to represent the cost data. These data contained Dutch costs from a healthcare perspective. The distributions are shown in Table 1.

Table 1 Distributions used for survivals and costs at time 0, distributions were fitted to figures published by Jansen (2005).

	Voriconazole	Amphotericin B
Survival (life weeks)	weibull (93.8,0.52)	weibull (46.6,0.45)
Costs (Euro per 12 weeks)	gamma (575,57)	gamma (570,59)

We then used Monte Carlo simulation to simulate from the above distributions and computed the INB in each stage based on all iterations. Using the central limit theorem we assume normality for the INB being a sample statistic. Its mean and variance were then estimated for each stage using the results of the simulation, for a range of willingness to pay values. Estimates of the INB values and their standard deviations for a range of willingness-to-pay thresholds at time 0 are reported in Table 2.

Willingness-to-pay (Euros per life year gained)	INB (s.d)
20,000	24,000 (180,000)
40,000	46,000 (360,000)
60,000	66,000 (550,000)
80,000	86,000 (740,000)

Table 2. Estimates of the initial INB for different willingness-to-pay thresholds

Table 2 shows that for all the willingness-to-pay thresholds the INB has a positive value in the beginning of the conditional reimbursement period. However, the relatively large standard deviations indicate high uncertainty in these initial estimates. As the willingness-to-pay rises, the initial INB becomes more favorable but uncertainty also increases.

Adoption proportion

Adoption proportion is a parameter that indicates what percentage of the patient population is using the new medication. Changing this proportion would change the number of patients in both arms of the registry, as well as the number of patients that are currently benefiting from a drug. Since the value of such a

parameter is often not available before the registry starts and is was not for our case study, we examined the results for a range of adoption proportions from 0% to 100%. When a single value for the adoption proportion was needed, i.e. when the effects of other parameters were to be tested, we used the equal 50%-50% rate for both arms.

Incidence and Mortality rates

There are several underlying conditions that cause the infection with aspergillus spp. Besides, hospitals differ in their background infection rate. Season as well as external circumstances (like building activities) also play an important role. To estimate incidence rates, we multiplied the incidence rate of the different underlying conditions by the incidence of aspergillus spp. within the condition and summed these. Data used to compute the incidence rates are listed in Table 3.

Table 3 Incidence of underlying conditions and aspergillus spp. infection among each category of conditions

Underlying condition	Incident cases in Netherlands (2008)	Infections caused by aspergillus spp. (%) (Pagano et al. 2006)	Number of infected cases	Reference for the incident cases
AML	625	7.11	44	(Netherlands Cancer Registry website (KNL))
ALL	226	3.87	9	(Netherlands Cancer Registry website (KNL))
CML	164	2.07	3	(Silver 2000)
CLL	625	0.36	2	(van den Broek et al. 2012)
NHL	3691	0.81	30	(Netherlands Cancer Registry website (KNL))
HD	447	0.32	1	(Netherlands Cancer Registry website (KNL))
MM	1048	0.27	3	(Netherlands Cancer Registry website (KNL))
Total	**6826**	**1.35**	**92**	

As shown in Table 3, in total 92 patients would be infected by aspergilluss spp. per year in The Netherlands. It should be mentioned that this is a very rough estimate in lack of appropriate data on incident cases of aspergilluss spp. We will get back to this issue in the sensitivity analysis and the discussion.

Mortality proportions were derived from the trial comparing voriconazole and amphotericin B (Herbrecht et al. 2002). 12-weeks mortality proportions were 29% and 42% for voriconazole and amphotericin B, respectively.

Prevalence

The uncertain nature of the aspergillus infection also affected the prevalence estimations. There is a widespread perception that the prevalence of aspergillosis has increased over the past decades. However, the evidence for such a claim has been usually gathered from longitudinal studies that are conducted in a single hospital, and may not be representative of all infected patients (Warnock 2007). In order to find an estimate for the infection prevalence, we assumed that the population is stable, i.e. the incidence rate does not change over time, and computed prevalence as a function of mean survival time and incidence rate (Keiding 1991). Survival time was measured in the registry as the effect measure for the drugs. Mean survival times reported in the Herbrecht trial (Herbrecht et al. 2002) were different for the voriconazole and amphotericin B arms. Therefore, we combined the survival times with the adoption proportion (*r*) to find the overall prevalence of the aspergilluss infections:

$$p = i \times r \times S_v + i \times (1 - r) \times S_a$$

Where S_v and S_a are the mean survival time of voriconazole and amphotericin B, respectively, and *i* is the number of incident cases per year. Considering a mean survival time of 174 life-weeks (3.35 years) for voriconazole and 116.1 life-weeks (2.23 years) for amphotericin B (Jansen et al. 2005), 92 incident cases per year and an adoption proportion of 50%, 256 patients would have experienced fungal infections at the beginning of the conditional reimbursement process.

Registry information

An important assumption regarding the registry information is that the registry is well-designed and that the data is corrected for selection bias. It is essential to have data with minimum bias in order to support the final decision; hence this is a relevant assumption also for the timing of the decision.

Characteristics of the registry, for instance the number of patients, sample variations, and costs of registering patients were important in the data update process. The number of patients who are registered in each stage is a function of the incidence rate and the proportion of the total population who are being registered. This would determine the number of new observations which will contribute to the update in information on INBs of the new drug. We assumed that 75% of the patients who received either voriconazole or amphotericin B would be registered and their data would be available for the authorities at the time of making the definite decision. The latter assumption was tested in sensitivity analysis.

Sample variation is also among the registry-related information that shows the level of homogeneity in registry data. High variations in data from the registry is indicative of a slow rate of information arrival, while smaller variations show that information is updated in a fast way. Since the decision is to be made ex-ante (i.e. before knowing the variation in the registry data), we ran the model for four different values of sample variations.

Costs of registry include 1) Fixed cost of setting up a registry (which is not dependent on the number of patients); 2) Cost of registering each patient (cost of screening, sending questionnaires, etc, these costs will be multiplied by the number of patients). Estimation of these costs was based on opinions from an expert involved in the Dutch patient registries for fungal infections.

Sensitivity analysis

We performed univariate sensitivity analysis (SA) for two types of input variables: first for those which are uncertain in this study but mostly certain to the decision maker. This category includes variables that can be assigned by the authorities. For instance, the willingness-to-pay threshold can be chosen by the

health policy decision makers. Drug costs can also be established after negotiations with the manufacturers: sometimes a discount in the price of the new drug is offered for the duration of the registry. Similarly, the discount rate is known to the policy authorities. Having access to the registry set up and costs information, decision makers can choose the best value for the parameters of this category. These parameters are listed in Table 4.

Table 4 Base case values of input parameters which are known to decision maker, along with values used in SA

Parameter	Basic value	Value in SA
Total population	16,000,000	------
Willingness-to-pay threshold (€/life year gained)	40,000	20,000
		60,000
		80,000
Mean treatment costs for the new drug (voriconazole) per patient per 12 weeks	€26,800	10^{-1} times the initial value
Proportion of the total population that is registered	75%	10%
		95%
Fixed cost of the registry	€ 125,000/ year for first 4 years € 75,000/ year for second 4 years	±50% of basic value
Variable cost of the registry per patient	€ 250	±50% of basic value
Discount rate	0%	3%

The second category includes uncertain variables that should be estimated from the registry data along with other sources. For instance, the proportion of patients who use voriconazole remains uncertain even after observing the number of patients in each group. Due to adverse effects patients might switch from one drug to another drug during the treatment time. Besides, the number of patients using each medication might vary over time. Therefore, the proportion using voriconazole would need to be estimated. Since no basic estimate is available for the adoption proportion, we examined the effects of this variable for 4 proportions, starting from

the extreme point of 0% (when voriconazole is not adopted yet) to the other extreme of 100% (when voriconazole is completely adopted).

Other variables can be found by observing the data from the registry are sample variations, costs of registry, and population registered. These parameters are mostly very case-specific, and it is difficult to find an approximation before implementing the registry process. Therefore, we examined the effects of changing the parameters to different values and reported the optimal time after the variations.

Effects of other variables, including the number of incidence cases, mortality proportions, and survivals were also tested using the univariate SA. Table 5 provides a list of these variables with values used for sensitivity analysis.

The value used in the SA for the incidence rate is an estimate from an informal registry in the North of The Netherlands (Raw data from University Medical Center Groningen. 2011). Base case overall survival was taken from the study by Jansen (Jansen et al. 2005), and the lower bounds and upper bounds reported in the same reference are used in sensitivity analyses. Other references for values used for sensitivity analysis are listed in Table 5.

Table 5 Base case values of input parameters which are uncertain to the decision maker, along with values used in SA

Parameter	Basic value	Value in SA	Source of value in SA
Adoption proportion	50%	0 %	---
		25%	---
		75%	---
		100%	---
Number of incidence cases per year	92	200	(Raw data from University Medical Center Groningen. 2011)
Ratio of sample standard deviation in the registry to deviation in the initial trial	10^{-1}	10^{-2}	---
		1	---
		10	---
		100	---
Survival time voriconazole	174.0 weeks	160.1	(Jansen et al. 2005)
		188.8	(Jansen et al. 2005)
Survival time amphotericin B	116.1 weeks	104.8	(Jansen et al. 2005)
		128.0	(Jansen et al. 2005)
Mortality proportion voriconazole (per 12 weeks)	0.29	0.38	(Denning et al. 2002)
Mortality proportion amphotericin B (per 12 weeks)	0.42	0.67	(Denning 1996)

RESULTS

Base case scenarios

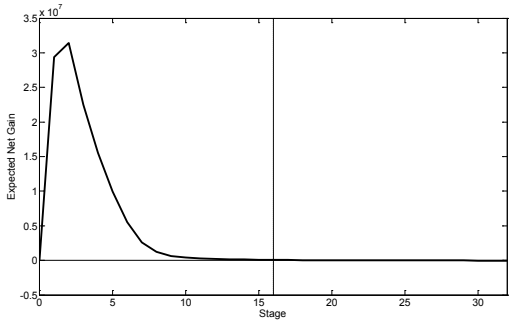
The base case ENG curve as well as the cumulative ENG for the scenario in which prevalent cases are registered as well as incident cases is depicted in Figure 1. As shown in Figure 1a, ENG is increasing during the first two stages and has a peak in the second stage. The prevalent cases are most likely registered in the first two stages, so the number of patients in the first stages is larger. This means a lot of new information is added in these stages. On the other hand, since the initial knowledge about the population is very uncertain, the information in the first stages considerably reduces uncertainty. Therefore, the curve peaks in the second stage. From the second stage onwards, the ENG curve constantly decreases over time and it reaches zero at stage number 21. However, since the negative numbers are very small, they are not distinguishable in the curve. Hence, the optimal time for making the decision in this case is stage 21 (~5 years). At this point, the cumulative gains are maximal (Figure 1b), so this would be the best time to make a definite decision. However, since the gains just before as well as the losses after the 21rd stage are small, the cumulative curve stays almost the same. So a definite decision between three and six years from the start would remain about optimal.

As shown in Figure 1, for the current strategy of 4 year delay (shown by the dashed line), ENG of a delay for one more period is above zero. However, the total area under the curve after the fourth year, which is the ENG of continuing the registry beyond 4 years, is limited. In this case, although the current policy of 4 years delay is sub optimal, it would not lead to huge losses.

Considering the scenario of registering only incident cases, the ENG curve shows a pattern quite similar to the first scenario shown in Figure 1, but the ENG in this case never approaches zero (figure not shown). Since the initial estimate of INB is relatively uncertain, the gain in information is large in the first few stages. When prevalent cases are not considered at all, the information is so scarce that observing the data from registry for only two stages would lead to a considerable gain. After the second stage, some basic information on the INB of voriconazole versus amphotericin B is gained, so the ENG starts to decrease. However, the gain

in information is so slow that the registry may continue for a long time without reaching an optimal point. The cumulative ENG keeps increasing when moving from one stage to the next. However, the increase gets very slow, with a pattern very similar to Figure 1b; hence again, any decision time from year 3 onwards would be about optimal.

a)



b)

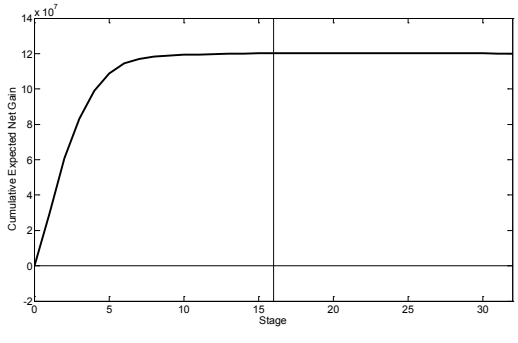


Figure 1 a) Expected Net Gains and b) Cumulative Expected Net Gains for the scenario of observing both prevalent and incident cases.

Sensitivity analysis

Results of one-way sensitivity analysis for the parameters known to the decision maker are listed in Table 6.

Table 6 indicates that higher willingness-to-pay thresholds will result in later time points of making the definite decision. When the system is paying more for an additional unit of health, a definite decision becomes more risky, hence it would be more appropriate to take time and observe sufficient data before deciding.

Reducing the treatment costs of the new drug (which mainly consist of the drug costs) does not affect the outcome. This shows that a discount (even as high as 90%) in the price of the drug during the registry period does not lead to shorter or longer optimal registry periods.

Sensitivity analysis on registry population shows that when only 10% of the population is covered by the registry, the information arrival is so slow that the cumulative ENG would not reach an optimal point within the first 8 years. Hence, the decision must be delayed to the latest possible time. On the other hand, a large registry that covers 95% of the population results in a somewhat earlier optimal time point of 19 stages comparing to the base case result of 21 stages.

As Table 6 indicates, the results are not much sensitive to the fixed registry costs, while the variable cost per patient has a significant effect. When registering patients is more costly, the point at which costs dominate gains is reached sooner. Low registry costs mean that observation can be continued longer before getting to the decision point. Discounting the costs and monetary gains does not change the optimal time either. As expected, since the time scale of the disease is short, discount rate does not cause much variation in results.

Table 6 Results of sensitivity analysis for the input parameters which are known to decision maker

Parameter	Basic value	Value in SA	Optimal stage
Total population	16,000,000	------	
Willingness-to-pay threshold (€/life year gained)	40,000	20,000	17
		60,000	26
		80,000	29
Mean treatment costs for the new drug (voriconazole) per patient per 12 weeks	€26,800	10^{-1} times the basic value	21
Proportion of the total population that is registered	75%	10%	>32
		95%	19
Fixed cost of the registry	€ 125,000/ year for first 4 years € 75,000/ year for second 4 years	+50% of basic value	21
		-50% of basic value	21
Variable cost of the registry per patient	€ 250	+50% of basic value	18
		-50% of basic value	29
Discount rate	0%	3%	21

Sensitivity analysis for the parameters of the second category (table 5) indicates that results are sensitive to the standard deviation of the registry data. When the standard deviation is as low as 10^{-2} times the initial distribution, data is more homogeneous and that makes the information update faster. Therefore, an optimal time point will be reached earlier (stage 17). High deviations would make the update slower which results in a decision time as late as possible.

For other parameters of this category, results are quite robust despite the high uncertainty in input parameters. Changing these parameters might change the overall ENG curves, but would keep the optimal time close to the base case result of 21 stages (i.e. about five years). According to results of the sensitivity analysis, the optimal time is not sensitive to the incidence rate and survival times. Similarly, it does not depend much on the prevalence of infections. Such robustness might be due to the fact that fungal infections are still rare in The Netherlands. Even doubling the number of patients per year would only increase the 12-weeks incidence rate from 1.4×10^{-6} to 3.0×10^{-6}, which is not a considerable change in absolute terms.

DISCUSSION

In this chapter we defined the optimal time of making a definite decision on reimbursement of voriconazole as the newest drug of choice for primary treatment of invasive aspergillosis. While the old regulations for expensive new inpatient drugs specified a maximum period of four years for conditional funding, we showed that in case of antifungal drugs the best time to make a definite decision is about the 5th year. However, making the decision at any point from the 3rd year on is also close to optimal for both scenarios considered: registering the prevalent and incident cases or registering only incident cases.

Results of this study have applicability outside the Dutch regulatory settings as well. Coverage with evidence development admission will in general get increasing attention in many jurisdictions (Carlson et al. 2010). Our results show how the decision makers could set the length of the period for evidence development such as to balance information gains against costs of further delay. Although the current study is aiming to optimize the Dutch reimbursement system which is mainly based

on real-life patient registry data, it can be modified for using other data sources in other jurisdictions. For instance, the framework proposed in this study can be applied to estimate the optimal follow-up time of a clinical trial during the coverage with evidence development process.

One advantage of the method used in the current study is that the duration of the research period is determined ex-ante. In this way, the time to make the definitive decision could be set at the beginning of the conditional reimbursement process, once a registry has started. The decision process could thus be completely transparent for the decision maker as well as the producer. In contrast, the problem of finding the optimal time for reimbursement decisions could be solved using methods that optimize the outcomes during the observation process. The real options approach (ROA) is an example of such methods, in which the optimal time is adjusted after observing the information in each stage (Dixit, Pindyck 1994). Another important and related method is the sequential analysis of trial data, in which the recruitment of the participants is continued until a convincing outcome is observed (Altman 1991).

Although the framework used in this study considered all relevant costs and gains affecting the optimal time of the conditional reimbursement period, a robust estimation depends both on including the right parameters and on having reliable estimates for these parameters. In our study, the high level of uncertainty in disease conditions implied that information on some parameters was very limited. Such uncertainties are due to the difficulties in diagnosis, different treatment options, and adverse effects which make physicians change the prescribed medicine during the treatment. In absence of reliable estimates for incidence and prevalence of fungal infections, we had to make assumptions which enabled us to find only rough estimates. The number of new incident cases was computed considering seven major categories of immunocompromised patients, while more categories of patients might be exposed to fungal infections due to reduced immunity. Another assumption was made to enable calculating prevalence as a function of incidence and disease duration. The assumption of stability might not be completely realistic, but it was the only possibility to estimate a base case for prevalence. Besides the limited data on the disease epidemiology, lack of registry data was a limitation that

would also be present in actual ex ante applications of this approach. Without the patient registry yet in place, estimates for costs and outcomes of the patient registry are bound to be uncertain. However, we have tried to find best estimates and have analysed the effects of variations in these parameters. As shown in sensitivity analysis, parameters related to incidence and prevalence did have a limited effect on the optimal time, as did the drug costs.

The proportion of voriconazole users is considered to be fixed over time in this study. However, the adoption proportion might be variable over time. It would usually take some time for a new drug to become widely available and accepted. Hence, an increase over time in usage of the new medication is reasonable. However, results of the sensitivity analysis illustrated that the optimal time in our case did not change much with the adoption proportion. Our model allows for evaluating changing proportions over time, but we do not expect the results to vary much for the current case.

To summarize, our application shows how to arrive at a reasonable estimate given the limited amount of data usually available at the time of a conditional decision.

From a broader point of view, the fact that the additional data is gathered using mainly registry data (i.e. observational study) might itself be a limitation for the process. Registries do not necessarily lend themselves for unbiased effectiveness estimate. For instance, when the new drug may be prescribed for a different patient population than the old one, the most efficient way of gathering additional data would be using an optimal portfolio of research combining different types of study designs including trials, epidemiological studies, surveys and patient registries (Conti, Claxton 2009). However, such an approach would be expensive and not very realistic when the new drug is already partially adopted. For the conditional reimbursement settings, registries are good solutions to gather additional data but only if they are well designed. Our study stresses the importance of carefully assessing registries in their early stages, thinking about their design and making sure that the registry data is going to provide the right evidence to support a final reimbursement decision.

The case presented in this study is a first step towards a flexible length for the

period of additional research for any new drug. The basic requirement for using this approach in its current form is that a well designed registry exists or starts in which new patients who are using either the new drug or the standard of care are being observed. Having the optimal re-evaluation time for any drug which is recently introduced, the decision makers would know when to look at the registry data and make the definite reimbursement decisions.

To conclude, due to the increasing number of new drugs and the limited health care resources, (conditional) reimbursement decisions urgently require improvement. In the current study we used the case of antifungal pharmaceuticals to show that decision making can be improved using a flexible approach which takes all the specifications of the drug and the disease into account. Thus the conditional reimbursement decisions can be modified for each new drug.

ACKNOWLEDGEMENTS

We would like to thank Professor Stephen Palmer for his helpful comments on an earlier version of this chapter. We also thank Professor Maureen Rutten-van Mölken, discussant of a previous version of this chapter at LolaHESG 2012 meeting in Almen, The Netherlands, and members of the audience who contributed to the discussion, for a number of helpful suggestions.

Financial support for this study was provided entirely by a grant from The Netherlands Organization for Health Research and Development (ZonMw).

The authors declare that they have no potential conflicts of interest.

REFERENCES

Raw data from University Medical Center Groningen. 2011.

Altman D. *Practical statistics for medical research.* Chapman & Hall: London 1991.

Ascioglu S, de Pauw B E, Donnelly J P,Collette L. Reliability of clinical research on invasive fungal infections: a systematic review of the literature. *Medical mycology: official publication of the International Society for Human and Animal Mycology* 2001;39(1):35-40.

Barnes R A. Early diagnosis of fungal infection in immunocompromised patients. *The Journal of antimicrobial chemotherapy* 2008; 61 Suppl 1i3-6.

Carlson J J, Sullivan S D, Garrison L P, Neumann P J, Veenstra D L. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy* 2010;96179–190.

Commissie Farmaceutische Hulp (CFH), 2007. *Farmacotherapeutisch rapport voriconazol (Vfend®) voor de primaire behandeling van bewezen of waarschijnlijke invasieve aspergillose.*

Conti S and Claxton K. Dimensions of Design Space: A Decision-Theoretic Approach to Optimal Research Design. *Med Decis Making* 2009;29(6):643-60.

Denning D W. Therapeutic outcome in invasive aspergillosis. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* 1996;23(3):608-615.

Denning D W, Ribaud P, Milpied N, Caillot D, Herbrecht R, Thiel E, Haas A, Ruhnke M,Lode H. Efficacy and safety of voriconazole in the treatment of acute invasive aspergillosis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2002;34(5):563-571.

Dixit A K and Pindyck R S. Investment under uncertainty. Princeton University Press: Princeton, NJ 1994.

Eckermann S and Willan A R.Expected value of information and decision making in HTA. *Health Econ* 2007;16195-209.

Espinel-Ingroff A In vitro fungicidal activities of voriconazole, itraconazole, and amphotericin B against opportunistic moniliaceous and dematiaceous fungi. *Journal of clinical microbiology* 2001;39(3):954-958.

Girois S B, Chapuis F, Decullier E,Revol B G. Adverse effects of antifungal therapies in invasive fungal infections: review and meta-analysis. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 2005;24(2):119-130.

Heinz W J and Einsele H. Caspofungin for treatment of invasive aspergillus infections. *Mycoses* 2008;51 Suppl 147-57.

Herbrecht R, Denning D W, Patterson T F, Bennett J E, Greene R E, Oestmann J W, Kern W V, Marr K A, Ribaud P, Lortholary O, Sylvester R, Rubin R H, Wingard J R, Stark P, Durand C, Caillot D, Thiel E, Chandrasekar P H, Hodges M R, Schlamm H T, Troke P F, de Pauw B,Invasive Fungal Infections Group of the European Organisation for Research and Treatment of Cancer and the Global Aspergillus Study Group. Voriconazole versus amphotericin B for primary therapy of invasive aspergillosis. *The New England journal of medicine* 2002;347(6):408-415.

Jansen J P, Meis J F, Blijlevens N M, van't Wout J W. Economic evaluation of voriconazole in the treatment of invasive aspergillosis in the Netherlands. *Current medical research and opinion* 2005;21(10):1535-1546.

Keiding N. Age-Specific Incidence and Prevalence: A Statistical Perspective. *Journal of the Royal Statistical Society.Series A (Statistics in Society)* 1991;154(3):pp. 371-412.

Lehrnbecher T, Frank C, Engels K, Kriener S, Groll A H, Schwabe D. Trends in the postmortem epidemiology of invasive fungal infections at a university hospital. *The Journal of infection* 2010;61(3):259-265.

Netherlands cancer registry website (KNL), *http://www.cijfersoverkanker.nl*.

Pagano L, Caira M, Candoni A, Offidani M, Fianchi L, Martino B, Pastore D, Picardi M, Bonini A, Chierichini A, Fanci R, Caramatti C, Invernizzi R, Mattei D, Mitra M E, Melillo L, Aversa F, Van Lint M T, Falcucci P, Valentini C G, Girmenia C, Nosari A. The epidemiology of fungal infections in patients with hematologic malignancies: the SEIFEM-2004 study. *Haematologica* 2006;91(8):1068-1075.

SILVER, R., 2000. Chronic Myeloid Leukemia. In: BAST RC JR, KUFE DW, POLLOCK RE, ET AL., EDITORS, ed, *Holland-Frei Cancer Medicine.* 5 edn. Hamilton (ON): BC Decker.

Stafinski T, McCabe C J,Menon D. Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *PharmacoEconomics* 2010;28(2):113-142.

Stafinski T, Menon D, Davis C,McCabe C. Role of centralized review processes for making reimbursement decisions on new health technologies in Europe. *ClinicoEconomics and outcomes research : CEOR* 2011;3117-186.

van den Bergh M F, Verweij P E, Voss A. Epidemiology of nosocomial fungal infections: invasive aspergillosis and the environment. *Diagnostic microbiology and infectious disease* 1999;34(3):221-227.

van den Broek E C, Kater A P, van de Schans S A, Karim-Kos H E, Janssen-Heijnen M L, Peters W G, Nooijen P T, Coebergh J W, Posthuma E F. Chronic Lymphocytic Leukaemia in the Netherlands: Trends in incidence, treatment and survival, 1989-2008. *European journal*

of cancer (Oxford, England : 1990) 2012;48(6):889-895.

van der Linden J W, Snelders E, Kampinga G A, Rijnders B J, Mattsson E, Debets-Ossenkopp Y J, Kuijper E J, Van Tiel F H, Melchers W J, Verweij P E. Clinical implications of azole resistance in Aspergillus fumigatus, The Netherlands, 2007-2009. *Emerging infectious diseases* 2011;17(10):1846-1854.

Walker S, Sculpher M, Claxton K and Palmer S, 2012. *Coverage with evidence development, only in research, risk sharing or patient access scheme? A framework for coverage decisions.* CHE Research Paper 77.

Warnock D W. Trends in the epidemiology of invasive fungal infections. *Nihon Ishinkin Gakkai zasshi = Japanese journal of medical mycology* 2007; 48(1):1-12.

APPENDIX 1

Maximum likelihood approach

When INB_k is the estimate of incremental net benefits using the simulation results at the end of stage *k* with variance of s_k^2 (*k=0,1,…i* and *k=0* represents the start point of conditional reimbursement) , the maximum likelihood estimator (MLE) for *INB* at the end of stage *i* is:

$$\hat{\theta}_i = \frac{\sum_{k=0}^{i} W_k INB_k}{\sum_{k=0}^{i} W_k} \quad with \quad W_k = \frac{1}{s_k^2} \qquad k = 0,1,...,i$$

And the distribution of the updated INB which uses the information of the entire registry period until time i will be as follows:

$$b_i \sim N(\hat{\theta}_i, \frac{1}{\sum_{k=0}^{i} W_k})$$

b_i reflects the information that will be available when the registry would be used for i stages.

APPENDIX 2

Value of information framework for registry data over time

If we call n_{Ai} and n_{Bi} the numbers of patients in the registry at the end of stage *i* receiving drugs *A* and *B* respectively, we have:

$$n_{Ai} = n_{A(i-1)} + (t \times r \times k \times P_r) - (t \times m_A \times n_{A(i-1)})$$

$$n_{Bi} = n_{B(i-1)} + (t \times (1-r) \times k \times P_r) - (t \times m_B \times n_{B(i-1)})$$

Where *t* is the duration of each stage, *r* is the proportion of patients using the new drug (A) in the country, *k* is the incidence rate of the disease in each stage, P_r is the population in the registry area, and m_A and m_B the mortality rates for patients using drugs *A* and *B* respectively. At *i=0* which is the conditional decision point we have:

$$n_{A0} = r \times P_r \times l$$
$$n_{B0} = (1-r) \times P_r \times l$$

Where *I* is the prevalence rate. If we call the total population of the country to be P_c (excluding the registry area), the total number of patients that can potentially benefit from the results of the research at stage *i* can be calculated similarly:

$$N_i = N_{i-1} + (t \times k \times P_c) - (t \times r \times m_A \times N_{i-1} + t \times (1-r) \times m_B \times N_{i-1})$$

At the end of each stage, if the mean of the distribution of b_i (appendix1) is positive ($\hat{\theta}_i$ >0), the gain of waiting one more stage versus adopting *A* is calculated. In this case the Expected value of Sample Information (EVSI) in stage *i* is:

$$EVSI_{Ai} = N_i \left\{ \int_{-\infty}^{0} -b\{f_{i-1}(b) - f_i(b)\} \; db \right\}$$

Where f_i is the INB probability density function of *A* versus *B* in stage *i*. If b_i is in favour of *B* ($\hat{\theta}_i \leq 0$), the gain of waiting for one more stage versus permanently adopting *B* would be calculated:

$$EVSI_{Bi} = N_i \left\{ \int_{0}^{\infty} b\{f_{i-1}(b) - f_i(b)\} \; db \right\}$$

Total costs of waiting (including opportunity losses) in stage *i* when $\hat{\theta}_i$ >0 would be computed as:

$$TC_{Ai} = n_{Bi} \times mean(b_i) + (n_{Ai} + n_{Bi}) \times C_{vi}$$

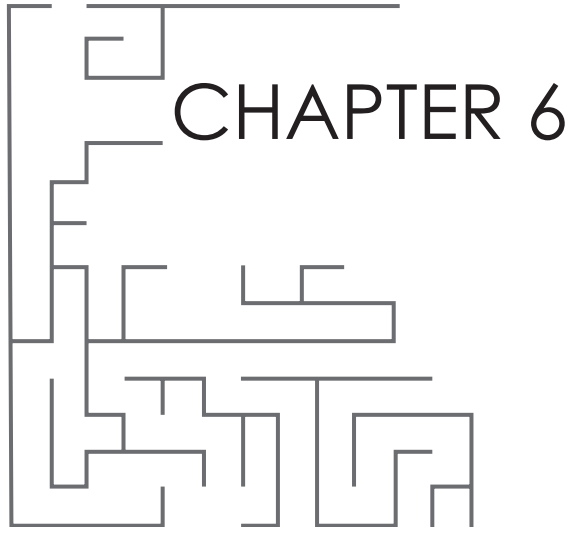
Where C_{vi} is the variable cost of being on the registry for each patient in stage *i*.
And total costs in stage *i* when $\hat{\theta}_i \leq 0$ would follow:

$$TC_{Bi} = n_{Ai} \times -mean(b_i) + (n_{Ai} + n_{Bi}) \times C_{vi}$$

Then the Expected Net Gain (ENG) of continuing registration rather than making a definite decision of the population in stage *i* would be:

$ENG_i = EVSI_{Ai} - TC_{Ai}$ *if* $\hat{\theta}_i > 0$

$ENG_i = EVSI_{Bi} - TC_{Bi}$ *if* $\hat{\theta}_i \leq 0$



Evaluation of patient registries supporting reimbursement decisions: The case of oxaliplatin for treatment of stage III colon cancer.

Leyla Mohseninejad
Chantal van Gils
Carin A. Uyl-de Groot
Erik Buskens
Talitha Feenstra

Submitted.

Abstract

Objectives: Access with evidence development has been established for expensive intramural drugs in the Netherlands. The procedure involves a 4-year period of conditional reimbursement. During this period, additional evidence has to be gathered usually through a patient registry. Given the costs and time involved in gathering the data, it is important to carefully evaluate the registry. This study aims to develop a model for regular evaluation of patient registries during an access with evidence development process and finding the optimal length of the registry period.

Methods: We use data from a recent registry in The Netherlands on oxaliplatin as a treatment option for stage III colon cancer. We add simulated further follow-up data to the empirical data available and apply value of information analysis to balance the gains of extending the period and amount of data gathering against the costs of registering patients.

Results: We show that given the assumptions on cohort size, follow-up time, and purpose of the registry, the current (partly simulated) registry was not very efficient. Notably, the observation period could have been stopped to make a definite reimbursement decision after 2 years rather than the fixed 4-year period.

Conclusions: Patient registries may be an efficient way to gather data on new medical treatments, but they need to be carefully designed and evaluated. For each purpose, data gathering can be tailored to make sure decisions are taken at the moment that sufficient data is available.

INTRODUCTION

The uncertainty in costs and effectiveness of new medical technologies makes it risky for decision makers to decide on reimbursement right after their market-approval by authorities. On the other hand, it is essential to keep pace with the rapid development of medical innovations. In recent years, the concept of "access with evidence development" (AED) has been introduced as a policy option to balance the careful evaluation of the new technologies with improvements in patient care by rapid access to technologies. Many developed countries have applied different forms of access with evidence development (Carbonneil et al., 2009; Carlsona et al., 2010; Mohr and Tunis, 2010; Trueman et al., 2010).

In the Netherlands, conditional reimbursement has been implemented as a way to ensure access while new evidence is being gathered. The current regulation is still under development, but previous regulation included a period of 4 years (Boer, 2012) in which the drug was reimbursed under the condition that during this period sufficient information on its cost-effectiveness would be obtained. After the collection of data in this period, the cost-effectiveness of the drug is to be reassessed in order to make a definite reimbursement decision.

Although randomized clinical trials are considered the gold standard for gathering data on drug (cost-) effectiveness, in conditional reimbursement settings patient registries are more attractive. That is because additional clinical trials are difficult to organize in the same population in which a new medication is already adopted and reimbursed, even conditionally. Due to lack of randomisation, registries have a serious risk of biased outcomes. This limitation may be managed to some extent by good design (Dugas et al., 2008) and using analysis techniques like propensity scores (Indurkhya et al., 2006). This is not the focus of the current manuscript though.

Advantages of registries are that they reflect daily practice more closely and can include a larger population since they do not require patients to agree to randomization for their treatment. Therefore, registries may be important sources of evidence (Gliklich and DeFilippo Mack, 2009). The current research focuses on patient registries as the source of additional data gathering. Setting up and

maintaining a registry is usually costly and time consuming and so it is important to evaluate its added value, both in advance and during the registry period. If a registry aiming to support a reimbursement decision does not produce, or is no longer producing information helpful for that decision, it is not worthwhile to continue gathering data and further delay the definite decision. In some cases, it might even be better to stop the registry and to use other sources of data (e.g. international trials) instead, unless other purposes warrant its continuation.

Our study gives an example of evaluating registries with the aim to support a reimbursement decision for the specific case of third line colon cancer treatment. We estimate the optimal duration of data gathering for a (partly hypothetical) registry. This duration could be very short or zero, actually indicating that the registry in its current form is not expected to add useful information for the decision concerning reimbursement. It might also be longer than actual follow-up, indicating that the follow-up time could have been longer to ensure a better decision.

Colon and rectal cancers are among the most common causes of death from cancer with 447,000 new cases and 215,000 deaths in Europe in 2012 (Ferlay et al., 2013). Since the 1990s, patients with stage III colon cancer were treated by adjuvant chemotherapy with 5-fluorouracil and leucovorin (5FU/LV) (Moertel et al., 1990). From 2005 onwards, National guidelines in the Netherlands have recommended the use of 6 months of treatment with 5FU/LV combined with oxaliplatin (FOLFOX) as the primary treatment option for stage III and possibly high-risk stage II colon cancer patients. As an alternative, the use of capecitabine combined with oxaliplatin (CAPOX) was also supported by the Dutch association for Medical Oncology (NVMO) (van Gils et al., 2012).

Treatment costs with oxaliplatin are quite high; hence the majority (80%) of oxaliplatin costs were reimbursed to hospitals in The Netherlands as of 2006 while a registry was initiated to provide additional information. This registry on stage III colon cancer patients was set up to collect additional evidence during the conditional reimbursement period. It has also provided information on guideline implementation in daily practice with respect to treatment choice, patient characteristics and dosage quantities (van Gils et al., 2012).

While the registry helps to gain new information and hence to reduce the

uncertainty in the decision, it also imposes costs consisting of set up costs and costs of registering patients. Moreover, during the conditional reimbursement period, no definite decision is taken and patients may receive suboptimal treatment as a result. In the methods and results section, it is explained how these may be balanced to evaluate the registry and find the best length for the period of additional data gathering. The discussion relates our findings to the actual data observation and decision process and considers how the process could be improved, changing the fixed 4-year period in the current regulations to a more flexible period.

METHODS

General approach

Using a healthcare perspective, we assumed that the definite reimbursement decision would get informed by the distribution of the Incremental Net Benefits (INB). The INB was calculated as $\lambda \times (S_O - S_c) - (C_O - C_c)$, where λ represents the willingness-to-pay threshold per disease free life year (DFLY) gained, S_O is the estimate of the disease free life years when using oxaliplatin (FOLFOX or CAPOX), S_c is the estimate of the disease free life years in the control population (5FU/LV or capecitabine), and C_O and C_c are estimates of the total costs for both types of treatment. Since the study aims to evaluate a partly hypothetical registry and decide on its optimal length, we estimated INB at several stages. Given disease prevalence and incidence, annual re-evaluation was assumed. Patients included in the registry were diagnosed in the period 2005-2006; hence the end of each year between 2006 and 2012 could have been a decision point (in 2012, the T=4 decision was scheduled). Starting with an initial distribution for INB at t_0 (the start point of conditional reimbursement period), the distribution of INB was updated after each year using the registry data. Having the consecutive distributions of INB, we calculated the gains obtained from the additional information after each year. Balancing these gains against the costs of the registry enabled us to evaluate the registry and decide on the optimal time of making a decision.

Patient population

A registry recorded information on stage III colon cancer Dutch patients who were diagnosed in 2005 or 2006, and who received adjuvant chemotherapy. The data was gathered retrospectively during 2008-2009. The database includes 391 patients that satisfy the inclusion criteria (van Gils et al., 2012), of which 281 patients had been treated with oxaliplatin. (FOLFOX in 136 patients and CAPOX in145 patients). The remaining patients received capecitabine (93 patients) or 5FU/LV (17 patients). Follow-up time before a relapse or censoring was reported, and used to estimate disease free survival (DFS). Drug costs and follow-up costs were also registered (van Gils et al., 2012). Since the aim of our study was to illustrate registry evaluation, we used the data presently available, i.e., not the final registry data.

Imputation of missing data

Some patients did not have a relapse and were censored at the end of the data collection period. However, for the purpose of this study we need to consider the case in which the data would have been gathered beyond 2008-2009. Therefore, from 2009 onwards, simulation was used to project the remainder of each patient's lifetime. The simulation was based on Weibull distributions for disease free life days fitted to the available patient data. Using conditional survivals, the expected future life expectancies were computed for all patients (see appendix 1). Medical costs were observed for the period 2005-2008 and used to simulate treatment costs and follow-up costs for the remaining years. We used constant costs per day for treatment phase and a gamma distribution on the proportion of total costs in each time interval during the follow-up phase, based on opinions of the involved experts (appendix 2 provides more details). This resulted in a partly empirical, partly simulated database covering the periods 2005-2009 and 2010-2012, containing information on all patients diagnosed in 2005-2006.

Survivals, costs and INB

Since the problem to address is the evaluation of the available registry data at potential points of making a definite decision, we looked at the partly hypothetical

data at the end of each year as if there would be no more information available after that date. This mimics how the procedure could be done prospectively for a new decision.

For each year, we filtered the (partly simulated) DFS to find the patients who had started treatment before the end of that year. If the patient experienced no event before the end of the year, the patient was censored. The costs of the censored patients were assumed to be the costs observed up to the censor date (appendix 2). This resulted in 7 different datasets containing the data observed up to the end of each year. These were analyzed to find their overall mean and standard errors of survivals and costs.

We assumed a λ of 60,000 €/DFLY gained (≈82 €/disease free life day), and changed this in the sensitivity analysis. The INB at the end of each year *i* is:

$$INB_i = (\lambda / 365) \times \left[E_i(S_O) - E_i(S_c)\right] - \left[E_i(C_o) - E_i(C_c)\right]$$

Where $E_i(X)$ shows the mean of parameter *X* at the end of the year *i*. Assuming independency between costs per day and the DFS time, the squared standard error of INB_i then is[1]:

$$s.e^2_{INBi} = (\lambda / 365)^2 \times \left[s.e_i(S_O)^2 + s.e_i(S_C)^2\right] + \left[s.e_i(C_O)^2 + s.e_i(C_C)^2\right]$$

This can be calculated for each year, using the number of patients in the registry at the end of each year (n_i).

Prior distribution of INB

The distribution for *INB* at t_0 (the start of the conditional reimbursement period) reflects the information available when the original decision to set up the registry was made. Using the MOSAIC trial (André et al., 2004), we found a mean and standard error for DFS months in each arm. Cost estimates used by decision makers at t_0 were obtained from consultations with experts involved in the registry. We conservatively assumed that the additional costs of oxaliplatin per patient per

[1] Given that for any random variable X, we have: $s.e^2(X) = std^2(X)/n$

treatment at t_0 had a uniform distribution with parameters (€0, €25000). Follow-up costs were equal for treatment with or without oxaliplatin as an initial estimate. The estimated INB and its standard error at t_0 was then calculated as:

$$INB_0 = (\lambda / 365) \times \left[E_0(additional\ \ survival)\right] - \left[E_0(additional\ \ \cos ts)\right]$$

And

$$s.e^2_{INB0} = (\lambda / 365)^2 \times (s.e_0(additional\ survival))^2 + (s.e_0(additional\ \ \cos ts))^2$$

INB updates

Having a prior distribution for INB and expressions for its observation in each stage, we used a Bayesian approach to find the updated INB values after each stage.

For notational simplicity, it is convenient to re-express the standard errors as the precision:

$$\tau_i \equiv \frac{1}{n_i \times s.e^2{}_{INBi}} \qquad i = 0,1,2,...,T$$

Where n_0 is the number of patients in the initial trial (2246 in total (André et al., 2004)) and n_i *(i=1,…T)* is the total number of patients in the registry in year *i*. After observing INB_i, the updated distribution of *INB* with respect to INB_i would be as follows (Christensen et al., 2010):

$$INB \left| INB_i \right. \sim N(\frac{\tau_0}{\tau_0 + n_i\tau_i} INB_0 + \frac{n_i\tau_i}{\tau_0 + n_i\tau_i} INB_i, \frac{1}{\tau_0 + n_i\tau_i}) \qquad i = 1,2,...,T$$

Gains and costs of additional follow-up time

At each potential decision point, the decision maker has the option to stop getting observations from the registry and make the definite decision or to postpone the decision for one more year. The Expected Net Gains (ENG) of waiting for the past year were found as the gains of waiting minus costs of waiting.

Gains as EVSI

Gains of waiting for more evidence before making a decision were basically calculated as the reduction in opportunity losses (Eckermann and Willan, 2007, 2008). Making a definite reimbursement decision means that either oxaliplatin is going to be routinely prescribed together with 5FU/LV or capecitabine for all patients with stage III colon cancer in The Netherlands, or it would be completely removed from the list of reimbursed drugs. The Expected value of Sample Information (EVSI) expresses the added value of gathering more information before taking the definite decision. The EVSI at the end of stage *i* was computed as the reduction in the opportunity loss from the end point of stage *i-1* to the end point of stage i. The opportunity loss expresses the possible losses resulting from a wrong decision. When uncertainty in the INB distribution is low, the possibility of a wrong decision decreases and hence the opportunity losses also decrease. The reduction in opportunity loss could be found from the changes in the distribution of INB after each stage. The detailed formulations of finding the opportunity losses and EVSIs are in appendices 3 and 4.

Number of patients who benefit from the decision

The number of patients in the country that can potentially benefit from a well informed decision after each stage multiplied by the gain per patient gives the overall gain of continuation of the registry. The total number of patients at the end of year *i* is:

$$N_i = N_{i-1} + (k \times P) - (r \times m_{oxaliplatin} \times N_{i-1} + (1-r) \times m_{no\ oxaliplatin} \times N_{i-1})$$

Where *k* is the incidence rate, *P* is the total population of the country, *r* is the proportion of patients who are using oxaliplatin, and *m* is the mortality proportion. Simply, we started with N_0 which is the number of prevalent cases at t_0, then the new incident cases were added and the dead cases are omitted after each year. Since the registry had started in 2005-2006, we used the prevalence and incidence rates of 2006 to estimate the total number of patients benefiting from the decision at each stage (N_i) (table 1).

Table 1 Population parameters used in the model

Parameter	Value	Reference
Proportion of oxaliplatin users	0.7	The Dutch registry
Population of the country	16,000,000	---
Number of prevalent cases in 2006	8,300	
Number of incident cases in 2006	1,900	

Costs of waiting

Costs of waiting include fixed costs of setting up the registry, which take place at t_0, and the variable costs of observing patients recruited over time. Therefore, at the end of the first stage the total costs (TC) would be:

$$TC_1 = C_f + (n_{1\,O} + n_{1\,C}) \times C_v$$

Where C_f is the registry set-up cost, C_v is the incremental variable cost per patient of being on the registry per year, and $n_{1\,O}$ and $n_{1\,C}$ are the number of registered patients for the first year in each group. From the second year onwards we have:

$$TC_i = (n_{i\,O} + n_{i\,C}) \times C_v$$

Gains versus costs: Expected Net Gains

Trading off the gains against the costs, the Expected Net Gains (ENG) of delaying the decision for one more year at the end of each year can be found. If the value of ENG turned out positive, the decision maker would know that so far gains had been obtained from delaying the decision, and could postpone the decision for one more stage. If the value appeared negative, the decision maker could stop the observation process because further continuation of the registry only implied more costs, unless the registry is providing gains for other purposes. Regardless of the latter, this would be the best time for the decision concerning definite reimbursement, since further delay would add no value. This is a rule of thumb, based on the assumption that the cumulative ENG shows a single peak.

Sensitivity analysis

The parameters related to the data and the update in INB were considered to be determined by the case study at hand, being outcomes of the registry. Hence for the sensitivity analysis, we focused on the parameters which were chosen by assumption.

The willingness to pay threshold was varied between €20,000/DFLY and €100,000/DFLY.

The base case value for the initial distribution of INB was based on MOSAIC trial, which is an international multicounty study. We used a wide uninformative prior with the mean 0 to test how the results would change without any information available at t_0. Population statistics show that incidence and prevalence of stage III colon cancer are increasing. Hence we also examined the effect of using the latest available data (i.e. data in 2012) on incidence and prevalence in The Netherlands.

RESULTS

Imputation of missing data

The scale (a) parameter and the shape parameter (b) for the fitted Weibull distribution[2] are shown in table 2.

Table 2 Specifications of disease free survival data based on the Dutch registry.

	Survival days distribution
Oxaliplatin	Weibull (1290.9,3)
No oxaliplatin	Weibull (1430,3.4)

Simulation of future disease free life time shows that most patients would have had a relapse by the end of 2012.

Survivals, costs and INB

Table 3 presents the results for the survival calculations as well as the treatment costs at the end of each year, resulting in incremental net benefits (INB). These

[2] Weibull CDF of $1-e^{-(x/a)^b}$ as the probability of event up to time x.

figures are based on actual empirical data until 2008 and after that they are based on simulated data.

In the final years, due to a longer observation period, more disease free days are observed and hence the mean is larger. Obviously that does not imply that mortality is decreasing.

Prior distribution of INB

DFS times of each arm and the incremental survival based on the MOSAIC trial (André et al., 2004) are reported in table 4. Expected additional costs of oxaliplatin had been estimated to be €12,500 for a planned treatment of 6 months.

As a result, the INB at t_0 has the following distribution:

$$INB_0 \sim N(-7500,1370^2)$$

Table 3 Number of patients observed per arm, means and standard errors of disease free survival days, costs, and INB using the data observed up to the end of each year

	Number observed		Mean disease free survival days (s.e)		Mean costs in Euros(s.e)		INB(s.e)*
Year	Control	oxaliplatin	Control	oxaliplatin	Control	oxaliplatin	
2006	103	260	640(18)	590(13)	5800(400)	17900(700)	-20580(870)
2007	110	280	920(29)	840(20)	8300(490)	21000(660)	-25900(970)
2008	110	280	1110(43)	1020(28)	9800(600)	23000(730)	-28010(1200)
2009	110	280	1150(47)	1050(31)	10900(720)	24500(810)	-30650(1350)
2010	110	280	1170(49)	1050(31)	11100(740)	24600(810)	-32200(1380)
2011	110	280	1170(50)	1050(31)	11200(750)	24700(810)	-33210(1390)
2012	110	280	1170(50)	1050(31)	11200(760)	24700(810)	-33040(1390)

* For illustrative purposes INB values are based on the available registry data combined with simulated data. When aiming to make an actual reimbursement decision, corrected real world data must be considered.

Table 4 Specifications of estimates at the start point of the conditional reimbursement (t_0)

	Mean (std)	Reference
FL+Oxaliplatin survival months	30(9)	(André et al., 2004)
FL survival months	29(10)	(André et al., 2004)
Additional survival months gained by Oxaliplatin	1 (13)	(André et al., 2004)
Additional costs of Oxaliplatin	12,500(720)	expert opinion

Expected Net Gains of additional follow-up time

Monetary gains of waiting for each potential decision point are reported in table 5. Fixed cost of set-up were by assumption €10000 and the variable costs €200 per patient per year. Table 5 also reports the resulting values for ENG and its cumulative value after each year.

Table 5 Expected Net Gains of delaying the decision after each year (numbers are rounded)

Year	Expected gains	Total costs	ENG	Cumulative ENG
2006	0	81800	-81800	-81800
2007	0	77400	-77400	-159200
2008	0	77400	-77400	-236600
2009	0	77400	-77400	-314000
2010	0	77400	-77400	-391400
2011	0	77400	-77400	-468800
2012	0	77400	-77400	-546200

The gains of the registry when considering DFS and costs as outcomes never exceeded zero. The ENG quickly converged to a value of -77400. This means that the data used in this example did not resolve the uncertainties around INB nor reduced the risk of the decision. Our results indicate that the efficiency of reimbursement decisions based on registries may be improved. The current uncorrected and partly hypothetical data on DFS and costs as relevant outcomes indicate that the registry could better have been stopped after one year of observation.

Sensitivity analysis

Table 6 shows the results of the sensitivity analysis. Only a willingness to pay as high as 100,000 €/DFLY gained implies a decision risk high enough to result in an optimal time of observation of 2 years. Assuming a very uninformative initial distribution for INB with mean zero rather than negative, also results in an optimal time of two year. That is, in absence of information at the time of conditional reimbursement decision, only two years of data observation would suffice for making a definite decision. Results for the prevalence and incidence rates indicate that the actual change in epidemiology of the disease during the period 2006-2012

did not affect the optimal time. Therefore, under the specific characteristics of this illustrative study (without inflow of new patients into the registry), a maximum of 2 years of data observation appeared sufficient for making a definite decision on reimbursement of oxaliplatin.

Table 6 Sensitivity analysis

Parameter	Base case assumption/ value	Assumption/ value in sensitivity analysis	Optimal registry time in sensitivity analysis
Willingness to pay	60,000 €/DFLY*	20,000 €/DFLY	1 year
		40,000 €/ DFLY	1 year
		80,000 €/ DFLY	1 year
		100,000 €/ DFLY	2 years
Prior distribution	$\sim N(-7500,1370^2)$	$\sim N(0,10000^2)$	2 years
Prevalence proportion	0.052 % (2006)	0.065 % (2012)	1 year
Incidence proportion	0.012 %	0.014 % (2012)	1 year

*DFLY: disease free life year

DISCUSSION

In this study we provided an example of evaluating the use of registry data to support the access with evidence process for reimbursement of oxaliplatin for stage III colon cancer treatment. We illustrated that the data observation could have been stopped after one or at most two years. Our study reinforces that setting up and continuation of a registry requires regular careful assessment of its results versus the expected outcomes.

Several studies have used value of information analysis to find the optimal design of clinical trials (e.g.(Chen and Willan, 2013; Eckermann and Willan, 2007, , 2008; Willan, 2008; Willan and Kowgier, 2008; Willan and Pinto, 2006)). Bayesian techniques have also been widely used in sample size determination (Halpern et al., 2001; Kikuchi et al., 2008; Pezeshk and Gittins, 2002). In contrast to these studies, the current study evaluates the optimal time to extract specific data from a given registry with a known design.

The method proposed in the current study takes into account the value of information in the data and considers the decision to be taken based on the registry

to find its value. In that sense it differs from the usual registry design considerations (e.g. data sources, patient selection, comparison groups, sampling strategies) (Gliklich et al., 2010).

There are several considerations in setting up a registry. The first issue is that the aims which the registry is meant to support must be well defined beforehand. For example, the aim might be to gather information on implementation issues in daily care like actual treatment costs or survivals. It might also be supporting a better informed decision on effectiveness or other outcomes for patients. Very often registries are designed to inform more than one parameter. For instance, the data studied in the current chapter has been shown to be very helpful in comparing the guidelines to the daily use of the chemotherapy with respect to treatment choice (van Gils et al., 2012). However, our model uses one specific decision objective (INB), covering two parameters (disease free survival and costs). To deal with this limitation, the timing may be based on the main purpose (the most important goal), or optimal times could be found for all purposes and the longest of these can be taken.

Like any observational study, registry data is inevitably biased. While solutions for this exist, for the real world data used in the current study, patient heterogeneity turned out to be too large to allow for appropriate correction of confounding in the registry data. This resulted in problems in estimating incremental cost-effectiveness using the registry data only (Franken et al., 2013). As a solution, a recent study (van Gils et al., 2013) has combined the registry data with the data from the MOSAIC trial (André et al., 2004) and the long term follow-up data of the trial (André et al., 2009) to find the cost-effectiveness of oxaliplatin. Our current case study is intended to illustrate the approach and hence we did not explicitly deal with these biases in the data and just presented the uncorrected outcomes. In real world applications, proper corrections should be included. Alternatively for cases that do not allow correction, the data could be used to confirm outcomes for the intervention arm only, rather than INB. That would result in net benefits being updated. It should be noted that even if the registry data can be successfully corrected for bias, it might not yet be worthwhile to gather additional data from it. Initiation and continuation of a registry is costly and if the gain in information is small for the purpose(s) which

the registry is meant to achieve, it is better to stop the registry and avoid extra costs or not to start it at all. Other purposes, e.g. scientific interest into the course of disease might warrant further follow-up though.

The current illustrative study takes the time factor of the registry into account and allows finding the time point when adding more observations provides no more gain in supporting the final reimbursement decision.

The registry evaluation was modeled as a so called a wait-and-see process. In this approach, the data gathering is stopped once a low (negative) registry net gain value is observed. This implies that a decision concerning registry continuation or cessation will be taken a posteriori, while the registry has already started. Hence, an a priori clear idea of the duration of the registry would not be available. Generally one would strive to make the definite reimbursement decision right after the optimal length of the observation period, i.e., when the amount of information contained has been achieved and processed. This length will change for different drugs and conditions, and though the procedures followed would be clear, their timing may be indeterminate. This might cause inconvenience for the policy maker, registry researchers and producers applying for reimbursement. One way to avoid this problem might be to simulate all possible registry outcomes, find an estimate for the optimal registry length and announce it to all the parties involved in the beginning of the conditional reimbursement period. However, using an entirely simulated data set would increase the uncertainty in the results.

Finally, a simple rule of thumb (stop a registry after showing negative expected net gains) was proposed for finding the stopping time. More sophisticated modeling could be used, applying methods from real option theory (Dixit and Pindyck, 1994). However, usually an analytical solution can only be obtained by imposing strict assumptions on the distributions of the parameters involved that will not be met in practice. Especially interdependence of outcomes over time (as present in the current dataset) is a problem.

To conclude, patient registries should not be considered a standard recipe for all access with evidence development procedures. Rather, they require careful design and should be used in the proper population and for the proper period, answering the proper research question. Continuation of the registry to support a reimbursement

decision while it is generating little gain in information can cause losses; hence it is essential to track its gains from the start and regularly re-evaluate it.

Acknowledgements

The authors would like to thank Dr. Wilbert van den Hout, discussant of a previous version of this chapter at LolaHESG meeting in May 2013 for his precise review of the manuscript and the helpful comment. The members of the LolaHESG 2013 audience who contributed to the discussion are also thanked for a number of helpful suggestions. Furthermore, we would like to thank the Dutch colorectal cancer group researchers who undertook the CAIRO studies for their cooperation.

REFERENCES

Netherlands Cancer Registry, Available from: http://www.cijfersoverkanker.nl/ [Accessed 3 Aug 2012]

André T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, Tabah-Fisch I, Gramont Ad. Oxaliplatin, Fluorouracil, and Leucovorin as Adjuvant Treatment for Colon Cancer. *New Engl J Med* 2004; 350: 2343-2351.

André T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C, Bonetti A, Clingan P, Bridgewater J, Rivera F, Gramont Ad. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J Clin Oncol* 2009; 27(19): 3109-3116.

Boer A. Procedure voorwaardelijk toelating geneeskundige zorg. College voor zorgverzekeringen (CVZ):2012.

Carbonneil Ce, Quentin F, Lee-Robin SH. A common policy framework for evidence generation on promising health technologies. *Int J Technol Assess* 2009; 25(Supplement 2): 56–67.

Carlsona JJ, Sullivana SD, Garrisona LP, Neumannb PJ, Veenstraa DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy* 2010; 96: 179–190.

Chen MH, Willan AR. Determining optimal sample sizes for multistage adaptive randomized clinical trials from an industry perspective using value of information methods. *Clin Trials* 2013; 10: 54–62.

Christensen R, Johnson W, Branscum A, Hanson TE. *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*. CRC Press: USA;2010.

Dixit AK, Pindyck RS. *Investment under uncertainty*. Princeton University Press: Princeton;1994.

Dugas M, Lange M, Berdel WE, Müller-Tidow C. Workflow to improve patient recruitment for clinical trials within hospital information systems – a case-study. *Trials* 2008; 9: 2-7.

Eckermann S, Willan AR. Expected value of information and decision making in HTA. *Health Econ* 2007; 16: 195–209.

Eckermann S, Willan AR. The Option Value of Delay in Health Technology Assessment. *Med Decis Making* 2008; 28: 300–305.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh J, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013; 49(6): 1374-403.

Franken MG, van Gils CW, Gaultney JG, Delwel GO, Goettsch W, Huijgens PC, Steenhoek A, Punt CJ, Koopman M, Redekop WK, Uyl-de Groot CA. Practical feasibility of outcomes

research in oncology: lessons learned in assessing drug use and cost-effectiveness in The Netherlands. *Eur J Cancer* 2013; 49(1): 8-16.

Gliklich R, DeFilippo Mack C. Clinical Trials Vs Registries. *April 2009 Supplement to Applied Clinical Trials* 2009: 1-4.

Gliklich R, Dreyer N, editors. *Registries for Evaluating Patient Outcomes: A User's Guide, Chapter 3, Registry Design.* Rockville (MD): Agency for Healthcare Research and Quality 2010.

Halpern J, Brown BJ, Hornberger J. The sample size for a clinical trial: A Bayesian-decision theoretic approach. *Stat Med* 2001; 20(6): 841–858.

Indurkhya A, Mitra N, Schrag D. Using propensity scores to estimate the cost-effectiveness of medical therapies. *Stat Med* 2006; 26(9): 1561-1576.

Kikuchi T, Pezeshk H, Gittins J. A Bayesian cost-benefit approach to the determination of sample size in clinical trials. *Stat Med* 2008; 27(1): 68–82.

Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Goodman PJ, Ungerleider JS, Emerson WA, Tormey DC, Glick JH, Veeder MH, Mailliard JA. Levamisole and Fluorouracil for Adjuvant Therapy of Resected Colon Carcinoma. *New Engl J Med* 1990; 322: 352-358.

Mohr PE, Tunis SR. Access with Evidence Development, The US Experience. *Pharmacoeconomics* 2010; 28(2): 153-162.

Pezeshk H, Gittins J. A fully Bayesian approach to calculating sample sizes for clinical trials with binary response. *Drug Inf J* 2002; 36(1): 143–150.

Trueman P, Grainger DL, Downs KE. Coverage with Evidence Development: Applications and issues. *Int J Technol Assess* 2010; 26(1): 79-85.

van Gils CW, Koopman M, Mol L, Redekop WK, Uyl-de Groot CA, Punt CJ. Adjuvant chemotherapy in stage III colon cancer: Guideline implementation, patterns of use and outcomes in daily practice in The Netherlands. *Acta Oncologica* 2012; 51: 57–64.

van Gils CWM, de Groot S, Redekop WK, Koopman M, Punt CJA, Uyl-de Groot CA. Real-World Cost-Effectiveness of Oxaliplatin in Stage III Colon Cancer: A Synthesis of Clinical Trial and Daily Practice Evidence. *Pharmacoeconomics* 2013: [Epub ahead of print].

Willan A. Optimal sample size determinations from an industry perspective based on the expected value of information. *Clin Trials* 2008; 5(6): 587–594.

Willan A, Kowgier M. Determining optimal sample sizes for multi-stage randomized clinical trials using value of information methods. *Clin Trials* 2008; 5(4): 289–300.

Willan A, Pinto E. The expected value of information and optimal clinical trial design. *Stat Med* 2006; 24(21): 1791–1806.

APPENDIX 1

The expected future lifetime

The probability of relapse at or before age t_1+t_2, given disease free survival until age t_1:

$$\Pr(t \le t_1 + t_2 \,|\, t > t_1) = \frac{pr(t_1 < t \le t_1 + t_2)}{p(t > t_1)} = \frac{F(t_1 + t_2) - F(t_1)}{S(t_1)}.$$

Where *F(t)* is the cumulative relapse probability at time t, and *S(t)* is the survival function, defined as:

$$S(t) = \Pr(T > t) = 1 - F(t) = 1 - \int_0^t f(u)du$$

Where *f(u)* is the probability of relapse at time point *u*. Therefore, the probability density of future lifetime is:

$$\frac{d}{dt}\frac{F(t_1 + t_2) - F(t_1)}{S(t_1)} = \frac{f(t_1 + t_2)}{S(t_1)}$$

And the expected future lifetime is calculated as:

$$\frac{1}{S(t_1)}\int_0^\infty uf(u + t_1)du = \frac{1}{S(t_1)}\int_{t_1}^\infty S(u)du$$

APPENDIX 2

The expected future costs and the costs per year

Medical costs took place over two different phases. First, the treatment phase during which drug and administration costs incurred. Since these costs were completely observed for all patients, there was no need to simulate them for the future simulated disease free life years (appendix 1).

In the follow-up phase, which starts after the treatment phase, costs are highest in the first two years. As follow-up continues, these costs decline and in the fifth year after treatment they are on average lower than €50 per patient per year. Hence, we fitted a gamma distribution, which describes the proportion of follow-up costs in each time interval after the start of the follow-up phase. The proportion of costs in the first *w* days of follow-up, *P(w),* is represented by the following notation:

P(w)=Gamma CDF(w, 1.5,400).

Where CDF stands for the cumulative distribution function. Using this proportion and the follow-up costs observed up to the data observation point, we estimated the costs related to the future simulated life expectancy of the patients.

To find the costs per year, we added the costs observed up to the end of each year. As an example, assume a patient has the pattern shown in figure 1.

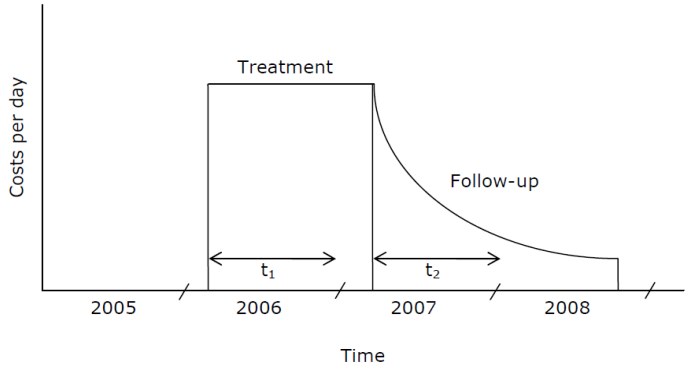


Figure 1. A hypothetical patient's treatment and follow-up costs

The costs observed by the end of 2006 (show here by C_{2006}) for the patient of figure 1 can be calculated as follows:

$$C_{2006} = \frac{t_1}{treatment\ \ time} \times treatment\ \ \cos ts$$

By the end of 2007, all treatment costs have been observed. Using the gamma distribution for the proportion of costs that fall in the first t_2 days of follow-up, for the patient of figure 1 the costs observed at the end of 2007 would be:

$$C_{2007} = treatment\ \ \cos ts + P(t_2) \times follow\ \ up\ \ \cos ts$$

At the end of 2008, all costs of this patient would have been observed so we simply have:

$$C_{2008} = treatment\ \ \cos ts + follow\ \ up\ \ \cos ts$$

In this way, for each patient in the database the costs for the end of each year were estimated and mean and standard deviation of costs over all patients were computed.

APPENDIX 3

Calculation of opportunity losses

Denote b_i as the estimate of incremental net benefits of A versus B at stage i with the mean $\hat{b}_i$. In case $\hat{b}_i$ >0, A would be adopted and the opportunity loss function of adopting A would be:

$L(i)= 0 \quad if \ b_i \geq 0$

$L(i) = - b_i \quad if \ b_i < 0$

In case $\hat{b}_i$ ≤0, B would be adopted and the expected opportunity loss of permanently adopting *B* would be as follows:

$L(i) = 0 \quad if \ b_i \leq 0$

$L(i) = b_i \quad if \ b_i > 0$

APPENDIX 4

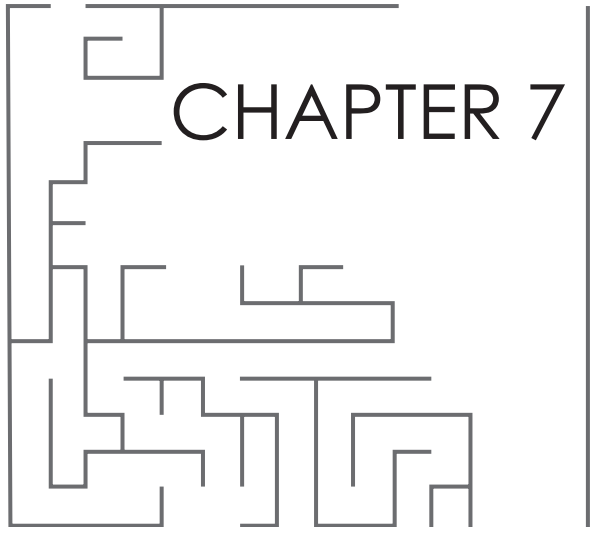
Expected Value of Sample Information

Denote f_i as the probability density function of incremental net benefits after stage *i*, *b* as the estimate of incremental net benefits of A versus B, and N_i is the total number of patients in the country who will benefit from the optimal choice at stage i. When mean *b* is positive, the gain of waiting one more stage versus adopting *A* is calculated:

$$EVSI_{Ai} = N_i\left[\int_{-\infty}^{0} -b\{f_{i-1}(b)\}db - \int_{-\infty}^{-0} -b\{f_i(b)\}db\right]$$

And if the mean *b* is negative, the gain of waiting for one more stage versus permanently adopting *B* would be calculated:

$$EVSI_{Bi} = N_i\left[\int_{0}^{\infty} b\{f_{i-1}(b)\}db - \int_{0}^{\infty} b\{f_i(b)\}db\right]$$



Summary and general discussion

In this thesis, several methods for handling uncertainty in economic evaluations and in policy decision making based on economic evaluations have been implemented, expanded and discussed. Deterministic and probabilistic sensitivity analyses together with different variations of value of information analysis were applied to analyze the uncertainty and its implications on decision making. Furthermore, optimal timing for making a decision was analyzed, given a situation of decision uncertainty and additional information becoming available over time. Results were illustrated using examples from different healthcare areas, namely prevention of depression by way of an e-health intervention, screening for presence of coeliac disease in patients with irritable bowel syndrome, third line treatment of colon cancer, and antifungal treatment for immune-compromised patients.

MAIN OUTCOMES

After an introduction to the concept of medical decision making and the role of uncertainty in **chapter 1**, we illustrated an elaborated case of cost-effectiveness modelling and analysis of uncertainty in **chapter 2**. We evaluated the cost-effectiveness of screening for coeliac disease in patients with diarrhoea/ mixed type irritable bowel syndrome (IBS) and showed that screening is a cost-effective way of improving quality of life/health for those patients. To get informed about the size and nature of uncertainty in the results, we performed scenario analysis, sensitivity analysis and value of information analysis. Results indicated that the uncertainty surrounding the cost-effectiveness estimate is limited, indicating that further research would not be worthwhile and deciding on implementation was possible based on the available information.

In **chapter 3** we examined the cost-effectiveness of opportunistic screening in combination with minimal contact psychotherapy (MCP), aiming to reduce the incidence of major depression. While the results showed that the program is cost-effective, analysis of uncertainty indicated a high risk in the decision of adopting the program; meaning that further research is required before making a final decision. We focused on analyzing the effect of perspective on the results of cost-

effectiveness and uncertainty analysis. In this way, we addressed an important methodological source of uncertainty, which is the selection of the perspective in the analyses. We showed that the selection of perspective has an important effect on the results of cost effectiveness and the value of information. Comparing the cost-effectiveness results from different perspectives showed that depending on the perspective chosen, the outcomes might significantly change and this might lead to an alternative decision. We estimated the expected value of perfect information for parameters to find the priorities for the further research, and showed that the selection of perspective would also affect the research priorities. When analyzing from a societal perspective, the priorities for future research were parameters related to productivity losses, while such parameters are not considered using a healthcare perspective.

Chapter 4 presented a framework for modeling and resolving uncertainty over time in an access with evidence development (AED) scheme. The model quantifies the gains and losses resulting from waiting for more observations from a patient registry before making a definite reimbursement decision on a conditionally reimbursed drug. Within the limitations of the model, it selects an optimal period of conditional reimbursement and additional evidence gathering. This can be used to find a best period in reality, while taking into account the model limitations and using sufficient sensitivity analyse to understand how the model outcomes depend on inputs used. That is, the model can be used to increase flexibility in timing of the access with evidence development process for different drugs and conditions. The model was developed to help decision makers find the time period needed for gaining additional evidence, whether the length of the period is to be defined ex-ante or ex-post. We illustrated our results using a hypothetical example for the ex-ante case, and discussed the advantages and points for improving the approach. We also explained how these methods relate to similar approaches like real options approach (ROA), sequential sampling and multistage trials.

In **chapter 5** we used the approach presented in chapter 4 and found the optimum time of making a definite reimbursement decision for voriconazole as the

newest drug of choice for primary treatment of invasive aspergillosis. We used an ex-ante variation of the model of chapter 4 in order to find the optimal length of data gathering process beforehand. Results indicated that the cumulative expected net gains (ENG) were maximized after 5 years of data observation. That means that after 5 years of gathering data from the registry, the uncertainty surrounding the cost-effectiveness estimate reaches an acceptable level: further delays in the definite decision would mean extra costs because of the opportunity losses and costs of the registry, while earlier points of making a decision would mean a higher risk in the decision due to the lack of evidence. However the increase in ENG after the 2nd year was not significant, indicating that a definite reimbursement decision in any point after 2 years of data observation was about optimal.

Chapter 6 applied an ex-post variant of the model developed in chapter 4, meaning that the length of the evidence development period is to be set as the data is being gathered (thus a wait-and-see process). This model was used to evaluate the use of registry data to support the access with evidence process for reimbursement of oxaliplatin for stage III colon cancer treatment. We showed that, if the registry only aims to gather information for supporting a reimbursement decision, the data observation could have been stopped after one or at most two years for this specific case. That means the registry only contributes to solving the uncertainties in the estimates of costs and effects of oxaliplatin for a limited period, and further delays imply losses.

Our results indicated that the approach for collecting additional data must be carefully chosen and constantly assessed.

IMPLICATIONS

As set out in the introduction, some methodological problems were addressed by the research reported in this thesis. We now turn back to these problems and see how the results in the thesis added to the state of the art. Furthermore, several chapters in this thesis were inspired by actual Dutch policy and clinical questions. This holds for chapters 2, 4, 5 and 6. Hence it is very relevant to discuss

the implications for Dutch policy. However, the international perspective is also relevant and the current section attempts to generalize results and discuss their international relevance.

Methodological and policy implications are not independent though. When the findings have a policy implication, methodology must be adapted to the new policy settings and when a new methodology implication is discovered, the policy settings might need to change to reflect the improvements in methods.

Policy and methodological implications concerning use of value of information analysis

Methodological implications

The use of value of information analysis is now described as one of the best practices for handling uncertainty by International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and Society for Medical Decision Making (SMDM) (Briggs et al., 2012). The methods used in this thesis are in line with the most recent changes in international policy settings. We used variations of the value of information analysis in all case studies of this thesis to address and present uncertainty in the most informative way. In chapter 2 we conducted value of information analysis alongside probabilistic sensitivity analysis to report complete evidence about the problem at hand (Mohseninejad et al., 2012). In chapter 3 we contributed to the applications of value of information analysis in policy making by showing that selection of perspective of the economic evaluation can significantly affect the expected value of information (EVPI) (Mohseninejad et al., 2013).

Implications for Dutch policy making

VOI analysis has been recommended in the Dutch guidelines as a method to identify the value of research and the critical parameters to be studied. Adding a VOI analysis allows assessing the reliability of the decisions based on the results of economic evaluations. This would add information compared to a (probabilistic) sensitivity analysis alone, since it can answer the question of whether conducting

additional research would be worthwhile and if so, what sort of research would be required. For instance, in chapter 2 of this thesis we evaluated the cost effectiveness of a screening program for irritable bowel syndrome (IBS) and we showed that the consequences of a wrong decision are marginal in monetary terms, which can be derived from the low value of additional research (Mohseninejad et al., 2012). Such results lead to a rather fast implementation of the recommended screening strategy in guidelines in The Netherlands (Van der Horst et al., 2012). In contrast, value of information analysis in chapter 3 showed that for the Dutch population and settings, further research is needed before making a decision on implementing the depression prevention program (Mohseninejad et al., 2013). Given that the societal perspective is most relevant for Dutch decision making concerning this type of programs, VOI research must also be conducted from a societal perspective.

Implications from an international perspective

After the first quantification of the expected value of obtaining more information about the intervention under study by Claxton (Claxton, 1999), interest in VOI arose, ultimately leading to the implementation of the method in the National Institute for Health and Clinical Excellence (NICE)'s technology appraisal procedure (Claxton et al., 2002). However, the growth in the application of VOI analyses was not as fast as the increase in the number of methodological studies in the beginning. In recent years, VOI analyses have been included in the policy procedures for technology appraisal in several countries and have started to be actually applied in reimbursement procedures. This change in the policy settings has lead to an increasing attention towards applying VOI in case studies of health technology assessment. As a result, the number of VOI case studies has reached the number of published methodological papers in 2010. (Steuten et al., 2013). However, even when included in guidelines, actual reimbursement decision making might not yet rely on VOI results (Claxton and Sculpher, 2006). We come back to this point and address the policy challenges regarding VOI analysis in the next section.

The role of perspective is also under discussion in different jurisdictions (Claxton et al., 2010). While many other countries, like The Netherlands, do use a societal

perspective, most VOI work, by its origins in the UK, still is limited to a health care perspective. As a general implication for policy settings, it is essential to run the value of information analysis considering the same relevant perspective as the economic evaluation. An invalid perspective could lead to unrealistic importance attached to additional research. Besides, as discussed in chapter 3, it is also important to choose the right willingness-to-pay threshold for the VOI analysis from different perspectives.

Access with evidence development (AED)

In chapters 4,5, and 6 of this thesis we analyzed the uncertainty over time in the AED schemes and we discussed different related policy issues. In this section I describe some methodological and policy implications of the findings of these chapters.

Implications for the Dutch practise

The AED scheme is equivalent to what is called "conditional reimbursement" in the Netherlands. Conditional reimbursement was firstly applied to the expensive drugs in the period 2006-2011, when the expensive new inpatient drugs were conditionally funded for a maximum period of four years. Until early 2012, the temporary funding of the drug costs (in which hospitals were receiving additional funding of 80% of drug costs) was based on available effectiveness data, a prognosis of cost-effectiveness and budget impact, and a plan for additional research. After the four years period, the drug was reassessed based on the additional information gathered during the period of conditional reimbursement and a definite decision was made. From 2012 on, the expensive drugs are part of the basic health insurance, but their costs are reimbursed as an add-on to a diagnosis-related group (DRG). This allows hospitals to ask for reimbursement of their costs in addition to the normal drug costs included in the DRG. Once the drug is admitted as an add-on, the reimbursement authorities can evaluate it and consider conditional reimbursement based on the research questions that have to be answered during the data gathering period. In the new regulation, the length of the observation

period is yet unspecified. In lack of any further suggestions however, 4 years still seems applicable.

The increasing concern about the timing of the reimbursement decisions in the Dutch policy settings lead to a research proposal written by the authors of chapters 4, 5 and 6 of the current thesis. Financial support for the studies was provided entirely by a grant from The Netherlands Organization for Health Research and Development (ZonMw) (2010). Findings of the current thesis are considered for the new guidelines on outcomes research, as proposed by the Dutch healthcare insurance board (CVZ). The goal is to find the best way to implement the model developed in chapter 4 of this thesis and have it integrated in the decision, with respect to the real world examples given in chapters 5 and 6. We are currently dealing with the feasibility/implementation of the approach in the regulatory process, by having organized discussions with CVZ and the association for innovative medicines in The Netherlands (Nefarma).

Implications from an international perspective

Implementing performance based reimbursement procedures have been an increasing trend over the last years. Different AED schemes, varying in design and complexity, have been implemented in several countries over the past years. The length of the AED period, i.e. the time between the point at which a new drug/technology had been launched and the point at which a policy decision was announced or expected to be announced ranged between 2–11 years (Stafinski et al., 2010) in different countries. Some of the schemes were time restricted after which a reassessment would take place based on the additional data generated. For other schemes a specific length for the data gathering period was not set in the beginning, but the data was subjected to periodic reviews (Carlson et al., 2010). In Europe reimbursement systems, the length of the additional research period is often not specified. In some countries (Belgium, Czech Republic, Denmark, Finland, and France), a fixed time, varying from 1 to 5 years, is set for all pharmaceuticals. In a few systems (Scotland, Sweden, and UK) the review period varies with the pharmaceutical (Stafinski et al., 2011) and the time for definite decision depends

upon the availability of the new evidence. However, the latter systems do not contain a robust framework for estimating the time needed for additional research. This is the common limitation of all AED schemes: they lack a systematic method to tailor the timing of either the reassessment or the periodic reviews to the characteristics of the technology and condition under assessment. Relaxing this limitation is an important policy implication of the findings of this thesis.

We used value of information analysis to develop a method to optimize the length of the additional data observation period, whether it is to be determined before the launch date of a new product (chapters 4 and 5) or after the arrival of new information (chapter 6). In chapter 4, we also present a user interface to enable calculation of optimal observation time as a function of different drug and population parameters. By using this simple user interface decision makers are able to explore and understand the relationship between the time needed for different cases to report reliable outcomes and the specifications of the drug, the disease, and the population.

POLICY CHALLENGES AND LIMITATIONS

Several issues need to be considered before applying the findings of this thesis in policy procedures. Some of the issues have already been addressed in this thesis, while some others are currently under discussion. Below I list some major challenges and I describe possible solutions.

Challenges regarding willingness-to-pay and the perspective of the analyses

Almost all types of economic evaluation need a certain willingness-to-pay (WTP) threshold for analysis. Without having an idea of the monetary value of each additional unit of health, making a decision remains problematic. Besides affecting the probability of a right decision (chapters 2 and 3), WTP threshold affects the future research priorities (chapter 3), the gains of waiting for more evidence (chapter 4), and the optimal time of decision making (chapter 5). Therefore, the uncertainty in selection of the right WTP threshold, which is categorized under “methodological uncertainty”, is an important issue to consider in policy procedures. Although in

this thesis we have analyzed the sensitivity of all the different outcomes to the WTP threshold, still a clear decision would need a certain threshold value or a range of possible values. This is especially important to consider in the Dutch context, in which no explicit threshold has been defined. There are only a number of unofficial threshold values mentioned in the Dutch literature. For preventive interventions for instance, a threshold of 20,000 €/QALY has been mentioned (van den Berg et al., 2008). This threshold was firstly used in a health care perspective setting (Casparie et al., 1998). However, it has also been used when considering a societal perspective (van den Berg et al., 2008). Attempts to find the appropriate WTP thresholds considering different perspectives are ongoing (Bobinac et al., 2010, , 2012).

Selecting the right WTP threshold is not only a concern for the Dutch decision making context. In UK for instance, National Institute for Health and Care Excellence (NICE) has been reluctant to specify a certain cost effectiveness threshold to be used in the decision making in the past (Devlin and Parkin, 2004). However, for some time, the NHS is one of the few decision makers using explicit thresholds, with a range of £20,000 to £30,000 per QALY gained (McCabe et al., 2008). Having such a range for WTP threshold can greatly enhance the clarity in the decision. Therefore, investigating more on the suitable WTP threshold could significantly improve the Dutch decision making procedures.

A potential challenge regarding the selection of a WTP threshold is the inclusion of the perspective of the analysis. As discussed in chapter 3, considering the same threshold values when analyzing from different perspectives is not very logical. There have been elaborative discussions on interpretation of WTP thresholds when considering different perspectives (Claxton et al., 2010). The requirements in estimating the WTP threshold value as well as the large effect that this value might have on the results of the analyses makes it an important element of the policy making process. Therefore, more attention must be given to this issue within the technology appraisal procedures in the future, especially in The Netherlands.

Value of information analysis: is it worth the trouble?

Following the inclusion of the VOI analysis in guidelines for health economic evaluations, concern has arisen about whether the costs and complexity of such analyses are counterbalanced by the information they would add. Some argue that due to the analytic complexity of the VOI methods, they are not necessarily intuitive to the policy makers and funders of research in healthcare (Ramsey et al., 2008). Besides, it is not clear if prioritization of future research through VOI would increase the chance that the priority questions will actually be answered through the extra research, as VOI does not weigh feasibility (Myers et al., 2011). Such limitations have led some researchers to seek for less-demanding approaches to perform VOI analysis (Meltzer et al., 2011). Some others suggest that normal sensitivity analyses are sufficient to address the need for additional data collection and research priority setting (Corro Ramos et al., 2013).

Several studies have suggested ways to improve the utility of VOI approaches for decision makers. Myers et al. (2011) suggest identifying ways to compare the impact of different prioritization methods on the likelihood that further research would actually answer the priority questions, identifying the appropriate resources (including technical expertise), defining the timelines and the appropriate level of modeling complexity. Ramsey et al. (2008) suggest that the methods of analyses must be transparent, flexible to accommodate a variety of endpoints, informed with available data that are acceptable to stakeholders and easy to interpret. Claxton and Sculpher (2006) argue that the separation of the remits for research prioritisation decisions from adoption and reimbursement is a barrier in implementation of VOI analyses in policy environment. They conclude that persuading decision makers to invest in VOI methods would be difficult before changing these circumstances. Hence, while the number of applications is rapidly increasing (Steuten et al.), there is still a need to further improve its applicability in decision making.

Overall, VOI methods are more that just an additional analysis in the current policy procedures. They need careful investigation of their goals, transparency and the adaptability to the system before they may improve the decision making process.

Issues in timing of the reimbursement decisions

In chapter 4, we developed a model to estimate the optimal time for making a definite decision on a conditionally reimbursed drug. In chapters 5 and 6, we applied variations of the model presented in chapter 4. However, this model is a first step in addressing the problem of timing in reimbursement decisions. Certain issues must be taken into account before applying this model in policy procedures. Some challenges are described in the following sub-sections.

Ex-ante vs. ex-post

The first consideration in implementation of our model is whether the time for a final reimbursement decision is to be determined ex-ante (as in chapter 5) or ex-post (as in chapter 6). Both methods have certain limitations and strengths, which where elaborated in chapter 4. In short, the ex-ante method is limited because it imposes additional uncertainty to the decision about time, while it is strong because it gives an idea about the length of the conditional reimbursement before hand. In contrast, the ex-post (wait-and-see) approach uses more reliable data to obtain the optimal time for a definite decision, but it makes the decision maker wait for data before being able to estimate a decision time. It also requires an approach to approximate a solution, which might not be the global optimum.

It real world decision making process however, the decision would not be so black and white. There is the need to get the advantage of both methods, while trying to minimize their limitations. As suggested in chapter 4, a combination of the two methods might improve the information they provide. For instance, an ex-post model can be used to evaluate the consecutive evidence gathered in the observation period while updating the inputs needed to run an ex-ante model. When the estimate of the optimal time resulting from an ex-ante approach becomes sufficiently robust, this estimate can be announced to the stakeholders. In this way, possible situations where the new evidence conflicts with the a-priori information would also be handled.

Other strategies might also be possible, but they first need to be clearly discussed among different stakeholders in reimbursement decisions before they can be implemented.

Data update issues

A prerequisite for the AED schemes is the need for setting up criteria for additional research. Poor or uninformative data would not help solving the uncertainties in the final decision. While randomized controlled trials (RCTs) can be potential sources of evidence to support the reimbursement decisions, we do not include them in the model for optimal timing of the decisions (with the exception of *i=0*). The reasons for this exclusion were that 1) because the drug is already adopted, it is infeasible to recruit patients to RCTs, 2) there is no guarantee for existence of a relevant global RCT during the period of conditional reimbursement, and 3) including external RCTs for supporting the reimbursement decision needs translation of evidence across jurisdictions.

For the same reasons, many of the AED schemes carried out to date have involved observational studies (e.g. registries or some similar system of epidemiological data collection) (Claxton et al., 2011). That has several advantages: The data gathered in this way is more relevant for the decision making, as it reflects the real world practise for the patient population of the interest. Furthermore, measurements of effectiveness are gathered rather than efficacy, which is needed for a cost-effectiveness analysis. It is also more ethical, given that the drug is already available and all patients must have the chance to benefit from it.

It has been argued that data from such studies might not be sufficient for making a definite decision, considering examples of mistakes made in the past when relying on observational with poor quality to inform definitive decisions (Stafinski et al., 2010). Attempts have been made to improve the registries aiming to provide additional information during an AED scheme as well as the data analysis methods used on the results. Examples show that most limitations can be overcome (Schluessmann et al., 2009; Tunis and Whicher, 2009).

In line with concerns about the usefulness of registry information during an AED process, we proposed to use the value of information framework suggested in chapter 4 to assess the information provided by a registry besides fining the optimal time of the data observation. Our results underlined the importance of careful assessment of the data gathering process considering design, population and the

observation period. Using our framework in addition to the common methods for evaluation of the data quality (e.g. study design, data sources, patient selection, comparison groups, sampling strategies) (Gliklich et al., 2010) can be helpful to prevent losses resulting from poor or uninformative data collection. For instance, when the analysis results in an optimal time of $i^*=0$, that would mean the registry must be improved in order to inform a definite decision. On the other hand, if i^* is too large, it might imply that the registry is generating the evidence with a very slow rate and it could be improved to update the information faster.

As discussed in chapter 6, several issues must be considered before the patient registries can be of any help for the reimbursement decisions. It is essential to set the aims which the registry is meant to achieve before starting one, and to design the registry in line with those aims. While a registry might not be helpful for one purpose, it might be good for other purposes, and hence should be continued. The time aspect is also important: even a registry which is generating helpful evidence to support the reimbursement decision might not be worthwhile to continue for a long time.

Apart from the goal setting and the design of the registry, it must be considered that a patient registry might need some time before generating reliable evidence. During the period in which the registry is being initiated, the costs are rather high due to the initial set up costs, while the information is poor because of limited number of patients recruited. Hence, the decision makers must allow a certain period of time before judging the quality of the evidence. The design and the recruitment procedure of the registry can be supervised during this period.

Although we justify the exclusion of RCTs from our model, we do not imply that the RCTs are not important in real world policy making. Evidence provided by RCTs is often reliable as a result of careful design and patient recruitment; hence they are widely used to update information supporting reimbursement decisions. In the Netherlands however, the main source for data observation during the conditional reimbursement period is patient registries. Therefore, incorporating the evidence from RCTs to the observations from a patient registry is a challenge for the timing problem of the current Dutch technology appraisal procedures. Overcoming this challenge requires clear knowledge about possible future RCTs and

their transferability (Eckermann and Willan, 2009) to the population of interest. The situation might sometimes become very complex though. In general, it is advisable to perform meta-analysis studies at consecutive points of time when an ex-post approach is adopted. In this way, the reduction of uncertainty over time can be used as a proxy for future changes. When using an ex-ante approach, predicting future number of patients who will be recruited to the trials might help to simulate future outcomes and the trend in information.

IMPLICATIONS FOR FUTURE RESEARCH

Several sources of uncertainty were addressed and discussed in this thesis, using value of information methods as the basic methodological approach towards uncertainty. In the next sub-sections I explain how the methods can be improved in the future to deal with a wider range of policy issues.

Sources of uncertainty

While we worked on different sources of uncertainty in this thesis, we did not simultaneously include all the sources of uncertainty in the outcomes. That is because different sources of uncertainty have a distinct nature, which makes them difficult and sometimes infeasible to integrate. This is not a new concern though; some recent studies have tried nevertheless to find ways to quantify different uncertainty sources and integrate them in the parametric uncertainty analysis (Bilcke et al., 2011; Bojke et al., 2006; Price et al., 2011). Using the proposed methods to account for all sources of uncertainty in one integrated model can help to improve the uncertainty analysis in future, whether the problem is to find the risk in the decisions, prioritization of future research or timing of the decisions. However, these methods need further developments before they can address the uncertainty around all particular types of decision-analytic models.

Value of information methodology

Value of information methods have a great potential for improving healthcare research and policy in the future. The key point for future research on VOI methods

is to make them more consistent with the real world policy settings. That is, to make them more transparent, more flexible, easier to interpret and easier to apply. As discussed in the last section, VOI methods still need careful assessment before they can routinely be used to support the policy decisions. As the methods are being applied in health care policy making in different jurisdictions, their methodology must also improve in order to reflect all the complicated elements in medical decision making. Some of the improvement points were addressed in the last section.

Uncertainty and timing of the decisions

The model presented in chapter 4 was an initial step in finding the optimal time for making reimbursement decisions under uncertainty. This model can be improved in future, in order to present a more realistic policy situation. The current model for instance only includes a yes/no final decision regarding the time for making a definite decision. This can be extended to include other options like restricted admission, limitation of indication or cost sharing arrangements. Furthermore, the initial decision of admission to the conditional reimbursement can also be flexible in the future extensions of our model. Including randomized clinical trials in the information update and accounting for the registry set-up time can also possibly improve the validity of the model. Finally, value of information analysis for parameters (EVPPI) can be used alongside the model to find the research priorities in each stage of time.

CONCLUDING REMARKS

Uncertainty analysis is an essential requirement in every decision within health care policy procedures. No decision is free of uncertainty, and uncertainty is costly. On the other hand, new technologies are emerging every day while healthcare resources are limited. Therefore, underestimating the effect of uncertainty in decisions would on average impose extra costs to the healthcare system and society at large. Analyzing different types of uncertainty would give the health care stakeholders ideas about the probability of a wrong decision, the consequences of

a mistake in decision making, the priorities for research when aiming to solve the uncertainties, and the optimal period to wait before making a decision.

Theoretical improvements in the analysis of uncertainty must be in line with the policy requirements in medical decision making. Including all aspects of the real world decisions in the analyses would help the methods to adjust better to the policy decisions.

ACKNOWLEDGEMENTS

I would like to thank all the people and organizations who helped us to get a clearer view on the policy implications and the potential challenges of the findings of thesis. The present discussion benefited form the inputs of Gepke Delwel, (now at the Dutch Ministry of Health, Welfare and Sport, but at the Dutch Health Care Insurance Board (CVZ) when most of this research was performed) Wim Goettsch from CVZ, Maureen Rutten-van Mölken, from the Institute for Medical Technology Assessment (iMTA), Ties Hoomans from the University of Chicago, Wilbert van den Hout from Leiden University Medical Center, Bart Heeg and Maarten Treur from PharMerit International, and Stephen Palmer from the University of York.

Attendees of the presentations based on chapters of this thesis in different national and international conferences are also thanked for their valuable inputs. Some helpful discussions were held concerning the findings of this thesis during LolaHESG 2010, 2012, and 2013, INFORMS Annual Meeting 2012, iHEA 2011 and 2013.

REFERENCES

Real options to support decision making on reimbursement of expensive new inpatient drugs, 2010, Available from: http://www.zonmw.nl/nl/projecten/project-detail/real-options-to-support-decision-making-on-reimbursement-of-expensive-new-inpatient-drugs/samenvatting/ [Accessed 24-07-2013]

Bilcke J, Beutels P, Brisson M, Jit M. Accounting for Methodological, Structural, and Parameter Uncertainty in Decision- Analytic Models: A Practical Guide. *Medical Decision Making* 2011; 31: 675–692.

Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. Willingness to pay for a quality-adjusted life-year: the individual perspective. *Value in Health* 2010; 13(8): 1046-1055.

Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. Valuing QALY gains by applying a societal perspective. *Health Econ* 2012; 2012: [Epub ahead of print].

Bojke L, Claxton K, Palmer S, Sculpher M. Defining and characterising structural uncertainty in decision analytic models. CHE Research Paper 9. Centre for Health Economics, University of York:2006; York.

Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health* 2012; 15: 835 – 842.

Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy* 2010; 96: 179–190.

Casparie AF, Van Hout BA, Simoons ML. Guidelines and costs [Article in Dutch]. *Ned Tijdschr Geneeskd* 1998; 142(38): 2075-2077.

Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. *Health Econ* 1999; 8(3): 269-274.

Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, McKenna C, Soares M, Spackman E, Youn J. Uncertainty, Evidence and Irrecoverable Costs: Informing Approval, Pricing and Research Decisions for Health Technologies. CHE Research Paper 69. The University of York, Centre for Health Economics:2011; York.

Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). *Lancet* 2002; 360: 711–715.

Claxton K, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics* 2006; 24(11): 1055-1068.

Claxton K, Walker S, Palmer S, Sculpher MJ. Appropriate perspectives for health care decisions. *CHE Research Paper 54* 2010:

Corro Ramos I, Rutten-van Mo¨lken MPMH, Al MJ. The Role of Value-of-Information Analysis in a Health Care Research Priority Setting: A Theoretical Case Study. *Medical Decision Making* 2013; 33: 472–489.

Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ* 2004; 13(5): 437-452.

Eckermann S, Willan AR. Globally optimal trial design for local decision making. *Health Econ* 2009; 18(2): 203–216.

Gliklich R, Dreyer N, editors. *Registries for Evaluating Patient Outcomes: A User's Guide, Chapter 3, Registry Design.* Rockville (MD): Agency for Healthcare Research and Quality 2010.

McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 2008; 26(9): 733-744.

Meltzer DO, Hoomans T, Chung JW, Basu A. Minimal modeling approaches to value of information analysis for health research. *Medical Decision Making* 2011; 31(6): E1-E22.

Mohseninejad L, Feenstra T, van der Horst HE, Woutersen-Koch H, Buskens E. Targeted screening for Coeliac Disease among irritable bowel syndrome patients: analysis of cost-effectiveness and value of information. *European Journal of Health Economics* 2012; November 2012: [Epub ahead of print].

Mohseninejad L, van Baal PHM, van den Berg M, Buskens E, Feenstra T. Value of Information Analysis from a Societal Perspective: A Case Study in Prevention of Major Depression. *Value in Health* 2013; 16(4): 490-497.

Myers E, Sanders GD, Ravi D, Matchar D, Havrilesky L, Samsa G, Powers B, McBroom A, Musty M, Gray R. Evaluating the Potential Use of Modeling and Value-of-Information Analysis for Future Research Prioritization Within the Evidence-based Practice

Center Program. Methods Future Research Needs Report. Agency for Healthcare Research and Quality:2011; Rockville.

Price MJ, Welton NJ, Briggs AH, Ades AE. Model averaging in the presence of structural uncertainty about treatment effects: influence on treatment decision and expected value of information. *Value in Health* 2011; 14(2): 205-218.

Ramsey SD, Blough DK, Sullivan SD. A forensic evaluation of the national emphysema treatment trial using the expected value of information approach. *Medical care* 2008; 46(5): 542–548.

Schluessmann E, Diel P, Aghayev E, Zwieg T, Moulin P, Roder C. SWISSspine: a nationwide registry for health technology assessment of lumbar disc prostheses. *European Spine Journal* 2009; 18: 851-861.

Stafinski T, McCabe C, Menon D. Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *Pharmacoeconomics* 2010; 28(2): 113-142.

Stafinski T, Menon D, Davis C, McCabe C. Role of centralized review processes for making reimbursement decisions on new health technologies in Europe. *Journal of ClinicoEconomics and Outcomes Research* 2011; 3: 117-186.

Steuten L, van de Wetering G, Groothuis-Oudshoorn K, Rete`l V. A Systematic and Critical Review of the Evolving Methods and Applications of Value of Information in Academia and Practice. *Pharmacoeconomics* 2013; 31: 25–48.

Tunis S, Whicher D. The national oncologic PET registry: lessons learned for coverage with evidence development. *Journal of the American College of Radiology* 2009; 6: 360-365.

van den Berg M, de Wit GA, Vijgen SM, Busch MC, Schuit AJ. Cost-effectiveness of prevention: opportunities for public health policy in the Netherlands [article in Dutch]. *Ned Tijdschr Geneeskd* 2008; 152(23): 1329-1334.

van der Horst HE, De Wit NJ, Quartero AO, Muris JWM, Berger MY, Bijkerk CJ, Geijer RMM, Woutersen-Koch H. The NHG guideline Irritable Bowel Syndrome (IBS) (first revision of the NHG guideline Irritable Bowel Syndrome). *Huisarts Wet* 2012; 55(2): 204-209.

Nederlandse samenvatting

Dit proefschrift gaat over onzekerheid rond de resultaten van economische evaluaties in de gezondheidszorg. In de verschillende hoofdstukken worden diverse methodes om onzekerheid in kaart te brengen en om de consequenties ervan te analyseren toegepast en uitgebreid. Een belangrijke rol daarbij speelt de analyse van de gevolgen van onzekerheid voor besluitvorming. De centrale vraag daarbij is wat het risico is van een bepaalde beslissing en hoe daarmee om te gaan. Een ander belangrijk onderwerp is het juiste tijdstip voor het nemen van een beslissing, gegeven dat de mate van onzekerheid in de loop van de tijd verandert.

De theoretische concepten zijn toegepast op concrete voorbeelden uit uiteenlopende hoeken van de gezondheidszorg, namelijk preventie van depressie door een “e-health” behandeling, screening op Coeliakie bij mensen met een prikkelbare darm, behandeling van gemetastaseerde darmkanker en behandeling van schimmelinfecties bij mensen met een ernstig verminderde afweer.

Belangrijkste bevindingen

Hoofdstuk 1 bevat een inleiding op het proefschrift. Dit hoofdstuk start met een korte uiteenzetting over economische evaluatie met besliskundige modellen. Vervolgens komen de belangrijkste methodes om onzekerheid te analyseren aan de orde, en manieren om zo goed mogelijke besluiten te nemen, gegeven een situatie met onzekerheid. Daarnaast wordt kort de beleidsachtergrond geschetst. Een aantal onduidelijkheden en omissies vormen vervolgens de aanleiding tot de vraagstellingen in dit proefschrift: hoe om te gaan met “value of information” vanuit een maatschappelijk perspectief, en hoe het tijdstip van besluitvorming te bepalen voor her evaluatie bij een voorwaardelijke vergoeding.

Hoofdstuk 2 is een toepassing van een economische evaluatie op basis van een besliskundig model met uitgebreide onzekerheidsanalyse. Het gaat over Coeliakie bij mensen met een prikkelbare darm. Ter ondersteuning van de richtlijn “Prikkelbare darmsyndroom” zijn verschillende screeningsstrategieën geëvalueerd. Algemene screening is vergeleken met helemaal niet screenen en met uitsluiten

van de groep met obstipatie van de screening. Bij deze mensen is Coeliakie namelijk niet te verwachten. Gerichte screening bij mensen met diarree/gemixte vormen van prikkelbare darm syndroom blijkt het meest kosteneffectief. Gevoeligheidsanalyses, zowel univariaat als probabilistisch, zijn gebruikt om de onzekerheid in kaart te brengen. Vervolgens is “value of information” analyse gebruikt om na te gaan of extra onderzoek toegevoegde waarde zou hebben. De onzekerheid bleek beperkt, zodat de waarde van extra onderzoek om de onzekerheid te reduceren marginaal was. (Mohseninejad et al., 2012). Dit betekent dat het redelijk is om de huidige resultaten te gebruiken voor besluitvorming. De nieuwe richtlijn bevat de aanbeveling voor gerichte screening. (Van der Horst et al., 2012).

Hoofdstuk 3 gaat over de toepassing van “value of information” analyse vanuit maatschappelijk perspectief. Bij de evaluatie van opportunistische screening op beginnende depressie gevolgd door “minimal contact psychotherapy”, (een gedragsinterventie met behulp van e-health) om ontstaan van een depressie te voorkomen is de omvang van de besparingen op productiviteitskosten zowel relevant als erg onzeker. Bij een maatschappelijk perspectief tellen deze besparingen uitdrukkelijk mee, terwijl een gezondheidszorgperspectief ze negeert. De meerderheid van de “value of information” analyses is uitgevoerd vanuit dit gezondheidszorgperspectief.

In deze toepassing heeft het perspectief veel invloed op zowel de kosten-effectiviteit als de waarde van extra onderzoek. De prioriteiten voor onderzoek zijn ook afhankelijk van het gekozen perspectief. Vanuit maatschappelijk perspectief blijken dit vooral parameters te zijn die met de productiviteitskosten te maken hebben.

De discussie betreft issues voor het vergelijken van de resultaten vanuit verschillende perspectieven, zoals de gebruikte drempelwaarde. Die is niet zonder meer gelijk wat vergelijking bemoeilijkt.

Hoofdstuk 4 geeft een theoretisch kader voor de analyse van onzekerheid in de loop van de tijd, in het kader van voorwaardelijke vergoeding en zogeheten “access with evidence development” constructies. Dit type overeenkomsten waarbij

middelen voorlopig worden vergoed, op voorwaarde dat aanvullend onderzoek plaatsvindt, is in opkomst bij vergoedingsbesluiten. Het is de vraag wanneer een her-evaluatie het beste kan plaatsvinden. In de oorspronkelijke regeling voor dure intramurale geneesmiddelen stond een periode van 3 jaar voor her-evaluatie, welke later is verlengd tot 4 jaar. Het is echter de vraag of dit voor ieder middel de juiste periode is.

Door systematisch de opbrengsten en kosten te kwantificeren van langere observatie van gegevens uit een patiëntenregistratie en verder uitstel van besluitvorming kan het tijdstip gevonden worden waarop de additionele baten van verder uitstel niet meer opwegen tegen de kosten. Dit kan zowel ex-ante, door simulatie te gebruiken, als ex-post, door herhaalde analyse van empirische data. Gevoeligheidsanalyse geeft vervolgens inzicht in de factoren die dit tijdstip beïnvloeden. Met inachtneming van de beperkingen van een dergelijk model kan deze aanpak bijdragen aan meer flexibiliteit en maatwerk in de opzet van trajecten van voorwaardelijke vergoeding.

In de discussie relateren we onze aanpak aan verwante technieken uit de economie, namelijk “real options analysis” en uit de epidemiologie, namelijk sequential sampling en multistage trials.

Het onderzoek voor dit hoofdstuk en de twee volgende werd gefinancierd uit een subsidie van ZONMW, binnen het programma Dure Geneesmiddelen (2010). Mogelijk zijn de resultaten bruikbaar voor de nieuwe leidraad uitkomstenonderzoek van CVZ.

In het onderzoek dat ten grondslag lag aan **hoofdstuk 5** is de benadering uit hoofdstuk 4 toegepast op het antischimmelmiddel voriconazol voor behandeling van invasieve aspergillose. In dit hoofdstuk wordt de beste periode voor her-evaluatie ex-ante bepaald, met een simulatiemodel. Dit simuleert de uitkomsten van een patiëntenregistratie op basis van gegevens over kosten van patiëntenregistraties en epidemiologie van schimmelinfecties. Het model is geprogrammeerd in MATLAB en beschikbaar gemaakt als los te gebruiken applicatie.

De resultaten laten zien dat de cumulatieve verwachte netto baten van extra observaties maximaal zijn na 5 jaar. Dat wil zeggen, na 5 jaar data verzamelen is

de onzekerheid rondom een kosten-effectiviteitsschatting gebaseerd op deze data van een zodanig laag niveau dat verdere onderzoek niet opweegt tegen de kosten. Die kosten bestaan uit registratiekosten en uit opportuniteitskosten, omdat patiënten mogelijk een suboptimaal middel gebruiken. Echter, al na 2 jaar is een forse afvlakking van de netto baten curve te zien, die aanduidt dat ook besluiten op enig moment na 2 jaar zo goed als optimaal is.

Hoofdstuk 6 onderzoekt de toepassing van de benadering uit hoofdstuk 4 bij gebruik van ex-post besluitvorming. Analyse van de (ongecorrigeerde) empirische gegevens uit een bestaande patiëntenregistratie van darmkanker is gebruikt voor evaluatie van het middel oxaliplatin voor de derdelijns behandeling. Het onderzoek laat zien dat -puur vanuit het perspectief van besluitvorming over vergoeding van oxaliplatin, en op basis van ongecorrigeerde data- de registratie in zijn huidige vorm kon worden gestopt na een of hooguit twee jaar. Dat wil zeggen, na deze periode was de bijdrage van de data aan het verminderen van de onzekerheid in de besluitvorming beperkt en woog niet op tegen de kosten van verder uitstel.

Deze resultaten benadrukken het belang van een goede opzet van patiëntenregistraties die worden opgezet om vergoedingsbesluiten te ondersteunen. Uiteraard hebben dergelijke registraties ook andere doeleinden. Ook dan blijft het echter belangrijk om de juiste observatieduur van gegevens uit de registratie ten behoeve van de her-evaluatie te bepalen. In bredere zin onderstrepen de resultaten van deze toepassing een weloverwogen inzet van extra onderzoek en de juiste keuze van parameters voor aanvullend onderzoek en bijpassende studie opzet. In de uiteindelijke her-evaluatie bleken namelijk aanvullende resultaten uit internationale klinische trials minstens zo belangrijk als de resultaten uit de patiëntenregistratie.

Het proefschrift sluit af met een **Discussie (hoofdstuk 7)**, die de uitkomsten van de hoofdstukken relateert aan internationale literatuur en in hun beleidscontext plaatst.

Acknowledgements

Four years have passed since the first time I entered the Netherlands; four years with a lot of experiences, loads of new things to learn, and of course, ups and downs. Despite many rainy days, I have enjoyed many of the days I spent here. What I never forgot was that “it takes rain to make rainbows”. Well, here comes a piece of my rainbow: my PhD thesis book. At the end of this chapter of my life, I want to acknowledge all the people who were by my side through these years and helped me to get it done.

The first person who truly deserves my gratitude is Dr. Talitha Feenstra, my daily supervisor without whom I would have never achieved my goals. Talitha, you have the biggest role in the accomplishment of this thesis. You spent a lot of time to help me during my PhD, even when you did not really have to. Thank you for being always there for me. The second person is of course Prof. Erik Buskens, my promoter. Erik, I was always getting enthusiasm in the meetings that I had with you. I would like to thank you for your insightful helps, in both academic and non-academic issues that I faced during my PhD. Talitha and Erik, both of you were never like a “boss” to me, but rather like a close colleague. Thank you for supervising me while having trust in me. I believe that is quite a skill you two have. I am sorry to miss great colleagues like you, but I will never forget the lessons that I learned from you.

My sincere acknowledgement goes to Prof. Maarten Postma and Dr. Cornelis Boersma who supported me whenever I needed help during different projects. Maarten, you have always been motivating me, since the very first days of my PhD when I learned the basics of pharmaco-economics from you until the time I submitted my thesis to you as a member of my reading committee. I hope I can continue to learn from you in future collaborations. Cornelis, you played a major role in helping me to find my way from research to practice. Your advices made my work closer to what the real world needs every day, and in the end, you were the one who helped me to find my ideal job in between research and practice. Thank you for being such an inspiration and support.

Beside Prof. Postma, I would also like to state my sincere gratitude to Prof. Kluin-Nelemans and Prof. Uyl-de Groot who accepted to be in my reading committee and spent time to critically review my thesis manuscript.

I would like to warmly thank the members of the HTA group for being such nice and friendly colleagues. Paul, Douwe, Henk, Elise, Karin, Gimon, Qi, Han, Dennis, Gera, Pepijn and Lilian! I always enjoyed the official meetings that we had, along with the chats during lunch or different conferences. My first special thanks among the group members go to Douwe, who was the person that introduced me to Talitha and Erik for this project and helped me with the technical difficulties I had. Thank you Douwe! My second special acknowledgement among my colleagues goes to Lilian. Lilian, working close to you means a lot of nice talks, sympathy and laughter (people passing from the 4th floor corridor can testify that!). I am very pleased to have you as my paranymph. That means a lot to me. Thank you for all your support during these four years.

Anna and Kim, I would also like to thank you for being my officemates. You were always very open to me when I needed your help. I have enjoyed sharing the work space with you a lot.

I mentioned most of my former colleagues, and here I would like to thank my new colleagues. I have been lucky to start my next job in October 2013 at PharMerit, Rotterdam as a junior research consultant. Although I am a new member to the company, I was warmly treated by everyone and soon I felt like home. Thank you all PharMeriterians for accepting me among you.

Working environment and scientific support play an important role for one to succeed, but more essential than that in my opinion is the emotional support. I was lucky to get this support from many people in my life, and particularly in this 4-year period. The first person to acknowledge when it comes to emotional support is Nastaran, my closest friend since years ago. I had the chance to have her by my side when I moved to the Netherlands. Nana, you helped me through all the ups and downs since the time our flight landed in Schiphol until this day. Thank you my

dearest fellow, I never forget what a great friend you are to me. Then I would like to thank Pantea, my full time friend in Groningen. It was such a joy to share those evenings and weekends with you, and I am sure it will continue to be. Thank you for all the good memories that you made for me in Groningen, and thank you for helping me also as my paranymph. I would also like to state my warm gratitude to Solmaz, a dear friend, a good company, a brilliant physician, a transportation expert and a Dutch language assistant! Thank you Solmaz, without your friendship and support life could have been really difficult! A special thanks goes to Niloofar, my dear friend in the Department of Epidemiology. Niloofar, I am pleased that I had you close to me whenever I needed company, thank you! I also like to thank all the other friends that made my days in the Netherlands shiny, among them I must mention Reza, Ali, Shabnam, Babak, Vahid, Parisa and Mehdi. Last but not least, I would like to pass on my warm regards to my friends all over the world whose continued friendship never faded, despite being far away. Thank you my dearest Anahita, for all the good times that you made for me during my trip to the States and your trip to Europe. We will always stay close even though we live far from each other. Thank you Faeze, Meri, Metol, Matin, Aida, Homa, Mergoor and many other fellows whose friendship is unforgettable to me.

My very deep gratitude goes to my family; however, no words can really show my love and appreciation towards them. Here I give it a try: First of all, I have been the luckiest girl in the world for having a great father whose love and support was there for me in every step that I took in my life. Baba, I love you so much and I am proud that I have a father like you. You are the first motivation for success in my life. Thank you for being my father, thank you for giving me such an endless love that has enabled me to manage my life through all the difficulties so far. Then I wish to express my love and gratitude to my beloved mother who left us too early, but her memory always stays with me and keeps my heart warm. Sincere thanks to my stepmother, a great company and friend. Thank you for supporting me all these years. My sister Elaheh, I am so grateful for having you in my life. Your affection even from far away always reminds me that I am never alone. Thank you for being my sister and my close friend. My little sister Fatemeh, I am so proud of you, so proud

of how you found your way in your life, and so sorry for not being close to you when you needed me. Thank you for being my lovely little sister and mate. My stepsisters Elham and Ensi, I must acknowledge you for the kindness and support you always had for me, and especially in the last four years. Thank you my dear sisters.

I would also like to acknowledge my second family for the hospitality and care they always had for me. Thank you nonna, nonno, Daniela, Gabriele and Giuseppe for everything, I have really enjoyed all the times that I spent with you in Italy.

Many lovely people have been playing important roles in my life, and that is why I have a long acknowledgement (no regrets of course!). But there is one more person to appreciate, someone very important to me, someone who has brought light to my days and nights. My amore Carmine, I am so happy that I have you next to me, so happy that I can share my life with you. When you are with me nothing can scare me. You were the precious gift of Groningen to me! Thank you so much for your love and your care.

I wish everyone a merry Christmas and a happy new year (2014).

Leyla

Curriculum Vitae

Leyla Mohseninejad was born on November 14, 1984 in Mashhad, Iran. After completing high school studies (major mathematics and physics), she took the Iranian university entrance exam in 2003. Ranked among the top 1000 of more than 500,000 participants, she was qualified to enter Amirkabir University (Tehran Polytechnic), one of the most prestigious technical universities in Iran. She did her bachelor studies in Industrial Engineering (system analysis specialization) in Tehran Polytechnic and graduated in 2007. Since she was ranked among the top 10% students during bachelor studies, she was exempted from the master degree entrance exam and hence could start her master studies in Industrial Engineering in the same department right away. During her master, she worked part time as a researcher in the field of operation research for ISACO, an Iranian company for auto parts and after-sale services. For her master thesis, she developed a framework for evaluation of information systems and used the educational portal of Amirkabir University as a case study. In 2009, Leyla obtained her master degree and moved to The Netherlands to start a PhD project in health economics in University Medical Center Groningen. She studied the impact of uncertainty and methods of uncertainty analysis in medical decision making, which resulted in the current thesis. In October 2013, she started her new job in pharmaceutical consultancy at PharMerit, Rotterdam.

SHARE Publications

This thesis is published within the **Research Institute SHARE** of the Graduate School of Medical Sciences (embedded in the University Medical Center Groningen / University of Groningen).
Further information regarding the institute and its research can be obtained from our internetsite: www.rug.nl/share.

More recent theses can be found in the list below.

((co-) supervisors are between brackets)

2013

Siebelink MJ. *The child as a donor; a multidisciplinary approach*
(prof HBM van de Wiel, prof PF Roodbol)

Sidorenkov G. *Predictive value of treatment quality indicators on outcomes in patients with diabetes*
(prof FM Haaijer-Ruskamp, prof D de Zeeuw)

Vu DH. *Clinical pharmacology of tuberculosis drugs and tuberculosis control in developing world; the involvement of private pharmacy and the individualization of treatment using dried blood spot*
(prof JRBJ Brouwers, prof DRA Uges, prof VH Le, prof DH Nguyen, dr JWC Alffenaar)

Sijtsma A. *Physical activity and overweight in young children*
(prof PJJ Sauer, prof RP Stolk, dr E Corpeleijn)

Rosicova K. *Regional mortality in Slovakia: socioeconomic indicators and ethnicity*
(prof JW Groothoff, dr JP van Dijk, dr A Madarasova-Geckova)

Bobakova D. *Youth subcultures and problem behaviours in Slovakia: hip-hop, techno-scene, metal, punk,skinheads and Roma*
(prof SA Reijneveld, dr JP van Dijk, dr A Madarasova-Geckova)

Arends I. *Prevention of recurrent sickness absence in workers with common mental disorders*
(prof JJL van der Klink, prof U Bültmann)

Theunissen MHC. *The early detection of psychosocial problems in children aged 0 to 6 years by Dutch preventive child healthcare; professionals and their tools*
(prof SA Reijneveld, dr AGC Vogels)

Bragaru M. *Sports and amputation*
(prof JHB Geertzen, prog PU Dijkstra, dr R Dekker)

Broesamle TC. *Designing health care services using systems thinking; a theory, a method and their application in the Dutch community pharmacy*
(prof JJ de Gier, prof JJ van der Werf)

Jong J de. *Antibiotics use in children; pharmacoepidemiological, practical and environmental perpectives in the Netherlands*
(prof LTW de Jong-van den Berg, dr TW de Vries)

Rettke HG & Geschwinder HM. *Long-term outcomes in stroke rehabilitation patients and informal caregivers*
(prof WJA van den Heuvel)

Fortington LV. *Enabling the elderly person with lower limb amputation through surgery, rehabilitation and long term care*
(prof JHB Geertzen, prof PU Dijkstra, dr GM Rommers)

Lako IM. *Depressive symptoms in patients with schizophrenia; count symptoms that count*
(prof K Taxis, prof D Wiersma)

Arnardottir AH. *Regulatory benefit-risk assessment; different perspectives*
(prof FM Haaijer-Ruskamp, prof PA de Graeff, dr PGM Mol, SMJM Straus)

Meijer A. *The forest trough the trees; investigating depression in patients with cancer and patients with myocardial infarction using systematic reviews and meta-analytic techniques*
(prof P de Jonge, dr HJ Conradi, dr BD Thombs)

Kuchenbecker WKH. *Obesity and female infertility*

(prof JA Land, prof BHR Wolffenbuttel, dr A Hoek, dr H Groen)

Rozenbaum MH. *Costs and effectiveness of extended vaccination strategies against pertussis and pneumococcal disease*
(prof MJ Postma, prof E Hak)

Kingma EM. *Intelligence and functional somatic symptoms and syndromes*
(prof JGM Rosmalen, prof J Ormel, prof P de Jonge)

Kedde JH. *Sexual health of people with disability and chronic illness*
(prof HBM van de Wiel, prof WCM Weijmar Schultz)

Horst PGJ ter. *Clinical pharmacology of antidepressants during pregnancy*
(prof B Wilffert, prof LTW de Jong-van den Berg)

Sinha R. *Adjustment to amputation and artificial limg, and quality of life in lower limb amputees*
(prof WJA van den Heuvel, prof P Arokiasamy, dr JP van Dijk)

2012

Pechlivanoglou P. *Applying and extending mixed-effects models in health in health economics and outcomes research*
(prof MP Postma, prof JE Wieringa, dr HH Le)

Verboom CE. *Depression and role functioning; their relation during adolescence and adulthood*
(prof J Ormel, prof WA Nolen, prof BWJH Penninx, dr JJ Sijtsema)

Benka J. *Living with rheumatoid arthritis: do personal and social resources make a difference?*
(prof JW Groothoff, prof JJL van der Klink, dr JP van Dijk, dr I Rajnicova-Nagyova)

Kalina O. *Sexual risky behaviour among Slovak adolescents and young adults; social and psychological factors*
(prof SA Reijneveld, dr JP van Dijk, dr A Madarasova-Geckova)

Crijns HJMJ. *Drug related safety issues affecting pregnancy outcome and concerning risk minimisation measures; emphasis on pregnancy prevention programmes*
(prof LTW de Jong-van den Berg, dr SMJM Straus)

Vries HJ de. *Working with pain; sustainable work participation of workers with chronic non-specific musculoskeletal pain*
(prof JHB Geertzen, prof JW Groothoff, prof MF Reneman, dr S Brouwer)

Karsten J. *On the threshold of disorder; definition and course of subthreshold depression and subthreshold anxiety*
(prof WA Nolen, prof BWJH Penninx, dr CA Hartman)

Abma FI. *Work functioning: development and evaluation of a measurement tool*
(prof JJL van der Klink, prof U Bültmann)

Rodrigues HCML. *Who's the patient? Ethics in and around maternal-fetal surgery*
(prof PP van den Berg, prof M. Düwell)

Munster JM. *Q fever during pregnancy; lessons from the Dutch epidemic*
(prof E Hak, prof JG Aarnoudse, dr ACAP Leenders)

Holwerda N. *Adaptation to cancer from an attachment theoretical perspective*
(prof R Sanderman, prof MAG Sprangers, dr G Pool)

Salonna F. *Health inequalities among Slovak adolescents over time*
(prof SA Reijneveld, prof JW Groothoff, dr JP van Dijk, dr A Madarasova-Geckova)

Kolarčik P. *Self-reported health and health risky behaviour of Roma adolescents in Slovakia*
(prof SAReijneveld, dr JP van Dijk, dr A Madarasova-Geckova)

Schreuder JAH. *Managing sickness absence; leadership and sickness absence behavior in hospital care*
(prof JW Groothoff, prof JJL van der Klink, dr CAM Roelen)

Hoen PW. *New perspectives on depression and heart disease*
(prof P de Jonge, prof J Denollet)

For more 2012 and earlier SHARE-theses see our website.